

Trabajo Fin de Máster

Análisis de iniciativas open data en España. Un estudio sobre la apertura de datos sanitarios en portales autonómicos.

Analysis of open data initiatives in Spain. A study on the openness of health data in regional portals

Autora

M^a Ángeles Bertos Quesada

Director

Gonzalo Marco Cuenca

Facultad de Filosofía y Letras

2019

BERTOS QUESADA, M^a Ángeles

Análisis de iniciativas open data en España. Un estudio sobre la apertura de datos sanitarios en portales autonómicos / M^a Ángeles Bertos Quesada; director Gonzalo Marco Cuenca. – Zaragoza: Universidad de Zaragoza, 2019.

Trabajo Fin de Máster de Consultoría en Información y Comunicación Digital de la Universidad de Zaragoza, 2019.

100 p. ; 29 cm

1. Datos abiertos. 2. Portales de datos abiertos. 3. Datos abiertos en salud. 4. Open data. I. – Bertos Quesada, M^a Ángeles. II. Marco Cuenca, Gonzalo, dir. IV.- Análisis de iniciativas open data en España. Un estudio sobre la apertura de datos sanitarios en portales autonómicos.

Resumen

La implementación de un gobierno abierto por parte de las Administraciones Públicas ha fomentado el suministro de datos públicos tanto a nivel internacional como nacional. La apertura de estos datos ayuda a los ciudadanos, no solo a entender mejor las acciones que los gobiernos llevan a cabo en el desempeño de sus funciones, sino a realizar un seguimiento más exhaustivo de los mismos en el tiempo. Estos datos abiertos, cobran especial importancia si una vez publicados, son reutilizados para crear valor (ofrecer un beneficio a la sociedad) y fomentar la transparencia de las administraciones. Para que esto ocurra, es necesario que existan herramientas que faciliten su acceso; los portales de datos abiertos son los mecanismos que ponen a disposición de los usuarios estos datos y facilitan su acceso. El objetivo de este trabajo es analizar el nivel de apertura de datos de las comunidades autónomas España a través de sus portales de datos abiertos, centrando el análisis en el estado actual de publicación de datos abiertos en materia de salud.

Palabras Clave

Datos abiertos, datos abiertos en salud, portal de datos abiertos, administración pública, gobierno abierto, comunidades autónomas, generar valor, reutilización

Abstract

The implementation of an open government by the Public Administrations has encouraged the provision of public data both internationally and nationally. The opening of this data helps citizens, not only to better understand the actions that governments carry out in the performance of their functions, but to undertake a more thorough monitoring of them over time. These open data become especially important if, once published, they are reused to create value (offer a benefit to society) and promote the transparency of administrations. For this to happen, there is a need for tools that facilitate access; open data portals are the mechanisms that make this data available to users and facilitate their access. The objective of this work is to analyze the level of data openness of the autonomous regions of Spain through its open data portals, focusing the analysis on the current state of publication of open data on health.

Keywords

Open data, open health data, open data portals, public administration, open government, autonomous communities, value creation, reuse

ÍNDICE

1. INTRODUCCIÓN	11
1.1 JUSTIFICACIÓN Y PROBLEMA DE INVESTIGACIÓN	11
1.2 ESTADO DE LA CUESTIÓN	13
1.3 OBJETIVOS.....	17
1.3.1 OBJETIVO GENERAL.....	17
1.3.2 OBJETIVOS ESPECÍFICOS.....	17
1.4 METODOLOGÍA	18
2. DATOS ABIERTOS.....	21
2.1 DEFINICIÓN DE OPEN DATA O DATOS ABIERTOS	21
2.2 PRINCIPIOS DE LOS DATOS ABIERTOS	23
2.3 BENEFICIOS DE LOS DATOS ABIERTOS	25
2.4 MARCO JURÍDICO	26
2.4.1 MARCO JURÍDICO EN EUROPA.....	26
2.4.2 MARCO JURÍDICO EN ESPAÑA.....	28
2.5 CONDICIONES DE USO Y LICENCIAS	31
2.6 CALIDAD DE LOS DATOS ABIERTOS.....	35
2.6.1 MODELO DE CALIDAD: LAS CINCO ESTRELLAS DE BERNERS-LEE	36
2.6.2 NORMA 178301:2015.....	38
2.7 FORMATOS DE DATOS ABIERTOS	40
3. DATOS ABIERTOS EN SALUD	43
3.1 PANORAMA ACTUAL	43
3.2 PANORAMA NORMATIVO PARA EL ACCESO Y TRATAMIENTO DE DATOS EN SALUD ...	45
3.3 BENEFICIOS	46
3.4 OBSTÁCULOS EN LA IMPLANTACIÓN	46
4. PORTALES DE DATOS ABIERTOS.....	49
4.1 ¿QUÉ SON LOS PORTALES DE DATOS ABIERTOS?	49
4.2 PORTAL DE DATOS ABIERTOS DE LA UNIÓN EUROPEA.....	56
4.3 PORTAL DE DATOS ABIERTOS DE ESPAÑA.....	58
4.3.1 PROYECTO APORTA.....	58
4.3.2 CATÁLOGO DE INFORMACIÓN PÚBLICA DEL SECTOR PÚBLICO.....	60
4.3.3 MAPA DE INICIATIVAS DATOS.GOB.ES.....	61
4.4 PORTALES DE DATOS ABIERTOS DE LAS COMUNIDADES AUTÓNOMAS ESPAÑOLAS ...	62

5. RESULTADOS	66
5.1 INDICADORES SOBRE LOS PORTALES	66
5.2 INDICADORES DE CONJUNTO DE DATOS O DATASET	67
5.3 INDICADORES DE CALIDAD DE LOS CONJUNTOS DE DATOS EN SALUD	72
5.4 APLICACIÓN DE INDICADORES EXTRAIDOS DE LA NORMA UNE 178301:2015 A LOS 16 PORTALES DE DATOS ABIERTOS ANALIZADOS EN ESTE TRABAJO.	75
6. CONCLUSIONES	85
7. BIBLIOGRAFÍA	90
ANEXO I. EVOLUCIÓN DEL LINKED GOBERMENT DATA	98
ANEXO II. EJEMPLO DE DATASET CON LOS CAMPOS REQUERIDOS DE INFORMACIÓN POR LA NTI	100

Índice de Tablas

Tabla 1. Principios de los datos abiertos.....	24
Tabla 2. Beneficios de la apertura de datos abiertos	25
Tabla 3. Esquema del Modelo 5 Estrellas Berners-Lee	37
Tabla 4. Formatos Reutilizables utilizados en la apertura de datos.	41
Tabla 5. Formatos no reutilizables en materia de open data	42
Tabla 6. Clasificación de formatos según el modelo 5 Estrellas de Tim Berners-Lee	42
Tabla 7. Componentes básicos de un portal de datos abiertos.....	50
Tabla 8. Componentes recomendables de un portal de datos abiertos.	50
Tabla 9. Cuadro resumen: categorías determinadas en la NTI.....	55
Tabla 10. Información básica aportada por datos.gob.es para el portal de Andalucía	62
Tabla 11. Información básica aportada por datos.gob.es para el portal de Aragón	62
Tabla 12. Información básica aportada por datos.gob.es para el portal de Asturias.....	62
Tabla 13. Información básica aportada por datos.gob.es para el portal de las Islas Baleares	63
Tabla 14. Información básica aportada por datos.gob.es para el portal de las Islas Canarias ...	63
Tabla 15. Información básica aportada por datos.gob.es para el portal de Cantabria	63
Tabla 16. Información básica aportada por datos.gob.es para el portal de Castilla La Mancha	63
Tabla 17. Información básica aportada por datos.gob.es para el portal de Castilla y León	63
Tabla 18. Información básica aportada por datos.gob.es para el portal de Cataluña.....	64
Tabla 19. Información básica aportada por datos.gob.es para el portal de la Comunidad Valenciana	64
Tabla 20. Información básica aportada por datos.gob.es para el portal de Extremadura	64
Tabla 21. Información básica aportada por datos.gob.es para el portal de la Xunta de Galicia	64
Tabla 22. Información básica aportada por datos.gob.es para el portal de La Rioja	64
Tabla 23. Información básica aportada por datos.gob.es para el portal de la Comunidad de Madrid	65
Tabla 24. Información básica aportada por datos.gob.es para el portal de la Región de Murcia	65
Tabla 25. Información básica aportada por datos.gob.es para el portal de Navarra	65
Tabla 26. Información básica aportada por datos.gob.es para el portal del País Vasco	65
Tabla 27. Comparativa de datos publicados en salud	72
Tabla 28. Clasificación de formatos de datos abiertos según el modelo 5 estrellas Berners-Lee por CCAA	74
Tabla 29. Evolución del Linked Open Data.	99

Índice de Ilustraciones

Ilustración 1. Datos, información y conocimiento.....	21
Ilustración 2. Clasificación de licencias Creative Commons	34
Ilustración 3. Esquema 5 estrellas Tim Berners-Lee	37
Ilustración 4. Representación de los dominios de la Norma UNE 178301:2015 sobre Datos Abiertos	39
Ilustración 5. Esquema de formatos de Open Data	40
Ilustración 6. Pantalla principal del Portal de datos abiertos de la NHS de Reino Unido	44
Ilustración 7. Infografía sobre el programa público PADRIS.....	48
Ilustración 8. Captura de la página de inicio del portal de datos abiertos de la Unión Europea.....	56
Ilustración 9. Captura de pantalla de la página de inicio del portal de datos abiertos de España.	59
Ilustración 10. Líneas de actuación del Proyecto aporta.....	59
Ilustración 11. Captura de pantalla de las temáticas establecidas en el portal datos.gob.es....	60
Ilustración 12. Mapa de iniciativas en España	61
Ilustración 13. CCAA con Portal de Datos Abiertos	66
Ilustración 14. PDA organizados según la temática de NTI	67
Ilustración 15. Dataset publicados por CCAA.....	68
Ilustración 16. Conjuntos de datos publicados en Salud por CCAA.....	69
Ilustración 17. Porcentaje representativo de datos en salud en el total de datos publicados ..	69
Ilustración 18. Publicación de datos por temáticas NTI	70
Ilustración 19. Conjuntos de datos de salud por temas especializados	71
Ilustración 20. Desglose de la materia específica Estadísticas en submaterias.	71
Ilustración 21. Formatos utilizados en la publicación de datos de salud	73
Ilustración 22. Formatos en los que publican datos de salud las CCAA	75
Ilustración 23. Condiciones de uso y licencias en PDA	76
Ilustración 24. Número de descargas de conjuntos de datos	77
Ilustración 25. Portales de datos abiertos con catálogo de datos.....	78
Ilustración 26. Conjuntos de datos documentados en los PDA.....	78
Ilustración 27. Mecanismos de búsqueda de los PDA.....	79
Ilustración 28. Catálogos de los PDA	80
Ilustración 29. Disponibilidad de sistema de acceso API o SPARQL	81
Ilustración 30. Disponibilidad de canal de Colaboración y Participación	82
Ilustración 31. Canal de resolución de quejas y conflictos	82
Ilustración 32. Fomento de la reutilización de datos en los PDA	83
Ilustración 33. Puntuación de los PDA según los indicadores establecidos por la NTI	84

Lista de abreviaturas

AAPP	Administraciones públicas
CC	Creative Commons
CCAA	Comunidades Autónomas
CSV	Comma-separated values
GA	Gobierno abierto
HTML	HyperText Markup Language
JSON	JavaScript Object Notation
KML	Keyhole Markup Language
LOPD	Ley Orgánica de Protección de Carácter Personal
LRISP	Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público
LTAIPBG	Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno
NTI	Norma Técnica de Interoperabilidad
ODS	OpenDocument Spreadsheet
PDA	Portal de datos abiertos
PDF	Portable Document Format
RDF	Resource Description Framework
RISP	Reutilización de la Información del Sector Público
RSS	Really Simple Syndication
SHP	Shapefile
TXT	Textfile
UNE	Asociación Española de Normalización
WFS	Servicio de características web
WMS	Servicio de Mapas por la red
XML	Extensible Markup Language
XLS/ XLSX	Microsoft Office Excel
ZIP	Formato de comprensión

Agradecimientos

Agradecer en primer lugar a mi director de TFM Gonzalo Marco Cuenca, por su ayuda incansable, su buena predisposición a guiarme en el desarrollo de este trabajo, y por las palabras de ánimo dedicadas en los momentos más necesarios.

A mis compañeros de trabajo, por haberse convertido en mi segunda familia en Zaragoza, haberme soportado durante el desarrollo del curso académico hasta la obtención de este título y haberme ayudado siempre que lo he necesitado.

Y por último a mis seres queridos, quienes desde el principio de mi formación académica siempre me han animado y apoyado a conseguir todos mis objetivos.

1. INTRODUCCIÓN

1.1 JUSTIFICACIÓN Y PROBLEMA DE INVESTIGACIÓN

Los bajos niveles de transparencia existentes en todo el mundo conducen a una ineficiente utilización de los recursos, y limitan la confianza de los ciudadanos en las Administraciones Públicas (Curto, 2017). En los últimos años éstas han visto la necesidad que tienen las personas de acceder a cierta información hasta ahora no disponible. Por eso tienen la voluntad de avanzar hacia un gobierno más abierto, basado en los pilares que el presidente de los Estados Unidos, Barack Obama, formuló en 2009 (Obama, 2009).

Obama expuso por primera vez los principios en los que debía basarse un Gobierno Abierto (GA) asentados en: la transparencia, la colaboración y la participación. Tres conceptos, que fomentaron la implementación de proyectos de GA en muchas administraciones de todo el mundo. Este movimiento de GA pretende mejorar a su predecesor, el gobierno electrónico o *egovernment*, que simplemente trasladaba la prestación de servicios públicos al modo electrónico aportando únicamente ventajas para el ciudadano en cuanto a la oportunidad horaria y el ahorro de desplazamientos (Curto, 2017).

Pero un GA tiene unos objetivos mayores como el de ser capaz de entablar una conversación permanente con los ciudadanos, tener en cuenta sus opiniones, y tomar las decisiones en función de las solicitudes recibidas. En este proceso de apertura a la sociedad cobra gran importancia, puesto que es un elemento común a todas las iniciativas de GA, el suministro de la información pública para que todos los agentes interesados puedan hacerse un juicio de valor no sesgado. Esta divulgación de información, además de mejorar la rendición de cuentas de las Administraciones Públicas (AAPP) y, por ende, la transparencia gubernamental, ha originado una corriente a nivel mundial de suministro de datos gubernamentales abiertos u open data (OD).

Los datos abiertos son «datos», que se distribuyen para su acceso e incluyen una licencia que permite su reutilización. La unión entre administración pública y ciudadano para dar la mayor difusión posible a la información es el principal valor de los datos abiertos. Como reconoce Llinares (2010), *“el espíritu de la filosofía de open data es coger datos que en sí mismos no ofrecen ningún valor y dotarlos de significado para convertirlos en conocimiento válido y con valor para los demás”*. Pero tal y como recuerda Martín Delgado (2014) *“la liberación de datos responde en estos momentos más a un ejercicio de voluntarismo político de algunos dirigentes que han sido conscientes de la importancia de ello que a una consecuencia de exigibilidad jurídica”*.

En este sentido, los portales de datos abiertos (PDA) se constituyen como la herramienta básica y fundamental en apertura de datos. Un portal de datos abiertos es el artefacto tecnológico que permite catalogar y divulgar conjuntos de datos abiertos en formatos reutilizables para que puedan ser libremente procesados como materia prima en los procesos de participación y colaboración propios de los modelos de gobierno abierto. A diferencia de la clásica web institucional, primera herramienta empleada para la divulgación de información pública en la red, el PDA supone un único punto de acceso al catálogo de datos. Del mismo modo, a diferencia de los llamados portales de transparencia que han proliferado en los últimos años, los datos se ofrecen en formatos que permitan su agregación y tratamiento automatizado, superando la divulgación en PDF (García y Curto, 2019).

La liberación de datos supone un potencial económico ya que constituye una base esencial para muchos productos de información digital. Es imprescindible asegurar la reutilización eficiente de la información pública para poder aprovechar todo su potencial y desarrollar nuevos productos, servicios y mercados que logren un mayor desarrollo económico y mayor creación de puestos de trabajo.

Sin embargo, muchos portales no cumplen con estos propósitos y parecen centrarse en una mera publicación de datos como un medio de justificar su existencia. En este sentido, es necesario velar por la calidad de los datos que se suministran y por una organización eficaz para que puedan ser localizados de una forma normalizada y precisa. Todo ello con objeto de facilitar una reutilización adecuada y, como ya se ha comentado, poder aprovechar todo su valor y potencial.

1.2 ESTADO DE LA CUESTIÓN

Realmente se desconoce cuando surge el Open Data y la reutilización, pero a lo largo de la historia se ha podido observar ejemplos que siguen una misma filosofía.

Ya en 1797, el físico americano Joseph Henry, fue propulsor de exponer al resto de la comunidad científica sus avances tecnológicos. Las leyes de la ciencia y cualquier beneficio obtenido a partir de ellas es patrimonio de la humanidad y deben ser compartidas. De esta forma, sus inventos como el electroimán y otros inventos posteriores, los puso a disposición de quien quisiera utilizarlos sin realizar ninguna patente (Pintos, 2012).

En agosto de 2010, la *BBC History Magazine* publicó un artículo¹ en el que relataba cómo Florence Nightingale, enfermera británica durante la Guerra de Crimea (1854-1856) utilizó datos estadísticos proporcionados por el Ministerio de Defensa Británico para generar un gráfico que representaba los datos de las bajas británicas y sus causas. En este gráfico se pudo comprobar que el porcentaje de muertes en batalla era siete veces menor que las muertes en los hospitales. Este hecho propició que Nightingale consiguiera apoyos para mejorar las condiciones higiénicas, alimentarias y de equipamiento de los soldados reduciendo la tasa de fallecidos en un 75%. Este ejemplo, puede servir de orientación para comprobar que el concepto de reutilización de información no es un concepto nuevo (Pintos, 2012).

No obstante, Open data tal y como lo conocemos es un concepto reciente que surge en entornos académicos. En 1957, se celebró el Año Internacional de Geofísica, evento en el que participaron prácticamente todos los países del mundo. Entre los multitudinarios actos que se celebraron cabe destacar la creación por parte del *International Council for Science* de un conjunto de centros de datos a nivel mundial que permitía clasificar, catalogar, almacenar y distribuir información entre los distintos centros e investigadores (Pintos, 2012).

Respecto a la Administración Pública, todavía tuvo que pasar bastante tiempo para que los políticos se dieran cuenta del potencial de la información del sector público, su

¹ www.historyextra.com/issue/august-2010

posible reutilización y el valor que podía generar. No fue hasta 1998 cuando desde la Comunidad Europea se redactó el Libro Verde sobre Información del sector público en la Sociedad de la Información (1998). Este documento describe el potencial de la información del sector público y la necesidad de su explotación y reutilización de cara a desarrollar la industria de los contenidos digitales basada en la información pública (Pintos, 2012).

Con el fin impulsar su aplicación, se publicó en el año 2000 el estudio *Commercial exploitation of Europe's public sector information* de la Comisión Europea (Informe Pira), en él se hace una primera aproximación del valor de la información del sector público, señalando que representaba en el año 1999 casi 1% del producto nacional bruto de la Unión Europea. Este estudio desencadenó en la redacción e implantación de la Directiva 2003/98/UE referente a la reutilización de la información del sector público, y cuyo objetivo era la creación de un marco homogéneo que garantizara unas condiciones de equidad para todos los agentes reutilizadores en todos los países de los Estados Miembros (Pintos, 2012). Posteriormente fue modificada a la Directiva 2013/37/UE relativa a la reutilización de la información en el sector público. Esta nueva directiva establece políticas de apertura a la información, propiciando la disponibilidad y la reutilización de la información en el sector público con fines privados o comerciales, con restricciones mínimas o nulas de carácter jurídico, técnico y económico, buscando favorecer la circulación de la información no solo para los agentes económicos, sino también entre el público en general.

Esta directiva, también está relacionada con la Agenda Digital Europea, iniciativa emprendida por la Comisión Europea (2010)² que, a su vez, promueve la comunicación de la Comisión Europea sobre datos abiertos como motor para la innovación, el crecimiento y la gobernanza transparente. Este documento es también un importante referente para la armonización de las normas y prácticas nacionales en materia de reutilización, proporcionando considerables mejoras para la reutilización transfronteriza

² La Agenda Digital Europea forma parte de la estrategia “Europa 2020” puesta en marcha por la Comisión Europea y promueve entre sus acciones clave la apertura de datos públicos para su reutilización, la simplificación del sistema de licencias para el intercambio de contenidos y la implantación de normas de interoperabilidad. Para más información sobre la Agenda Digital Europea se puede consultar: <http://ec.europa.eu/digital-agenda/>

de los datos³. En este contexto, el portal de datos de la Unión Europea⁴ se abrió al público en una versión beta a finales de diciembre del año 2012 y se inaugura oficialmente en enero del año 2013 (Arquero y Marco, 2014).

En España, la primera línea de actuación nacional dedicada específicamente a impulsar la apertura de información del sector público fue el Plan Avanza. Posteriormente en 2007, se encomendó esa tarea de promoción a la Entidad Pública Empresarial Red.es y, pocos meses después, se aprobó la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, considerada realmente como la transposición al ordenamiento jurídico español de la directiva 2003/98/UE. Ya en el año 2009 nació el Proyecto Aporta⁵, un proyecto pionero de cara a fomentar el uso de la reutilización de información del sector público, y difundir su valor social y económico en los diferentes niveles de la administración, sector privado y ámbito académico de acuerdo a un modelo de gobernanza integrador. A posteriori, muchas comunidades autónomas (CCAA) comenzaron a desarrollar sus propios portales de datos abiertos.

Respecto al tipo de información que se puede reutilizar, de acuerdo a la Ley 37/2007, toda información generada por la Administración Pública es reutilizable, siempre y cuando no se encuentre excluida de forma expresa tal y como indica en su artículo 3. Adicionalmente, el Real Decreto 1495/2011 establece que la información del sector público es reutilizable sin necesidad de solicitar ningún tipo de autorización (siempre y cuando que la administración no tenga obligación de no liberar la información y proteger los intereses públicos y privados). Así mismo es responsabilidad de cualquier entidad pública con información reutilizable autorizar la reutilización de los documentos elaborados o custodiados por ellos. Los casos en los que la información no puede ser liberada son: información excluida por intereses públicos (para preservar el funcionamiento de las administraciones) e información excluida por intereses terceros

³ Muchos estados miembros de la Unión Europea han generado sus propias normas, derivadas de la Directiva 2003/98/CE, superando el nivel mínimo fijado en dicha Directiva, pero que han supuesto un obstáculo a la hora de la reutilización de los datos de forma transfronteriza, ya que para poder conocer el alcance del uso de los datos había que acudir a las diversas legislaciones nacionales.

⁴ Portal de datos abiertos de Europa <https://www.europeandataportal.eu>

⁵ Antiguamente www.aporta.com, que ha sido sustituido por datos.gob.es para adaptarse a la nueva normativa.

(preservar los derechos de terceras partes). Mientras que la información que si es susceptible de ser reutilizada es: información económica y comercial, geográfica, agrícola y pesquera, social, jurídica y legal, científica, cultural, política e información relativa a la gestión empresarial (Pintos, 2012).

Dorado López (2019) afirma que la automatización en la generación y puesta a disposición de los ciudadanos de la información susceptible a ser reutilizada, sigue siendo la acción más compleja de todas. Sin un proceso automatizado es muy arduo tener un portal de datos abiertos actualizado. La generación manual de la información, conversión de formatos, así como su subida al portal es una tarea muy tediosa y que sigue consumiendo muchos recursos hoy en día. Además, afirma que recopilar los datos de las distintas fuentes para subirlas al portal de datos abiertos en los diferentes formatos y manteniendo actualizados los metadatos asociados a cada conjunto de datos es un trabajo necesario para el éxito del proyecto pero que al requerir tanto esfuerzo puede hacer que la organización se relaje o desborde con respecto a su mantenimiento. Si los datos no están actualizados los proyectos de apertura están abocados al fracaso.

Tener un portal de datos abiertos hoy en día es sólo un primer paso que permitirá cambiar la percepción de la administración pública por parte de la ciudadanía, pero habrá que seguir yendo más allá, ya que la información en sí misma es excesiva y en algunos casos poco inteligible, y se debería trabajar más en mostrarla en formatos más intuitivos, de manera resumida incluso, para que sea sencillo extraer conclusiones por parte de quien la lee.

1.3 OBJETIVOS

1.3.1 OBJETIVO GENERAL

Este trabajo tiene como objetivo principal estudiar las iniciativas de gobierno abierto en materia de publicación de datos en las diferentes comunidades autónomas españolas a través de sus portales de datos abiertos y, en concreto, analizar el estado de publicación de los conjuntos de datos de salud en estos portales. Cabe señalar que se pretende conocer el estado actual de apertura de datos a nivel autonómico en nuestro país, ya que las AAPP autonómicas son las principales gestoras de la mayor parte de servicios de carácter social dirigidos a los ciudadanos. La investigación, por consiguiente, se propone analizar los portales de datos abiertos de las comunidades autónomas con dos fines principales; el primero, conocer el estado actual de publicación de datos abiertos y, segundo, realizar una comparativa más detallada sobre el estado de apertura de los datos abiertos en salud.

1.3.2 OBJETIVOS ESPECÍFICOS

- Comprender el concepto de apertura de datos abiertos y por qué es importante su publicación hoy en día.
- Localizar y estudiar los portales de datos abiertos de las CCAA para determinar el grado de dificultad de acceso a los datos abiertos publicados por las diferentes CCAA.
- Analizar y comparar los conjuntos de datos (o dataset) de los diferentes portales para conocer las principales temáticas de publicación de datos en España.
- Analizar los conjuntos de datos publicados por temáticas para verificar qué y cuánto se está publicando sobre *salud* con respecto al total de datos publicados.
- Determinar cómo es la calidad en relación a los formatos de publicación de los datos publicados por CCAA en salud.
- Detectar posibles carencias para realizar propuestas de mejora.

1.4 METODOLOGÍA

En primer lugar, se ha procedido a realizar un estudio de la literatura ya publicada con el objetivo de contextualizar el trabajo y conocer el punto de partida del mismo. Para ello, se realizó una serie de búsquedas bibliográficas en diferentes bases de datos, accesibles a través de la biblioteca de la Universidad de Zaragoza, como: Web of Science (WoS), Scopus o Dialnet. También se utilizó Google y Google Académico, para ampliar la literatura encontrada (en español o inglés). Además, con objeto de conocer la normativa vigente (leyes, decretos, normas) en materia de datos abiertos, se ha utilizado la base de datos de legislación del Boletín Oficial del Estado (BOE)⁶ y la base de datos de acceso al Derecho de la Unión Europea, EUR-Lex⁷. Otra fuente principal de información muy importante para la realización de este trabajo, ha sido el portal oficial de Datos abiertos del Gobierno de España⁸, quien nos ha servido de base para la conocer el panorama actual en dicha materia.

En segundo lugar, se definieron una serie de indicadores que han permitido dar respuesta al objetivo de esta investigación, y se procedió a acceder a cada uno de los portales de datos abiertos de las comunidades autónomas para realizar la extracción de los datos necesarios. Para dicha extracción, se utilizó el catálogo de datos⁹ existente en cada portal.

La metodología para la toma y análisis de los datos ha sido la siguiente:

Recogida de datos. La recogida de datos se realizó del 29 de julio al 4 de agosto de 2019. Se estableció extraer de cada portal la misma información para poder analizarla en una hoja de cálculo que permitiera realizar el análisis de forma más rápido y eficaz. Para acceder a los diferentes portales, se utilizó al portal oficial del Gobierno, datos.gob.es, a través del cual se accedió al mapa de iniciativas autonómicas en materia de datos

⁶ Disponible en: <https://www.boe.es/buscar/>

⁷ Disponible en: <https://eur-lex.europa.eu/>

⁸ Disponible en: <https://datos.gob.es/>

⁹ Catálogo de datos: sistema de información pública accesible vía web donde se almacenan y distribuyen datos y sus metadatos (UNE 178301:2015)

abiertos. Este mapa enlazaba a su vez con cada uno de los portales oficiales creados en cada comunidad autónoma.

Instrumento de recogida de datos. Los datos se han volcado en una hoja de cálculo en Excel, donde se han confeccionado una serie de tablas maestras con los datos recopilados. Los datos han sido revisados, clasificados y normalizados para su posterior análisis por medio de tablas dinámicas. Las tablas elaboradas han sido las siguientes:

- *“Información general”*: Se ha recopilado información sobre si una CCAA disponía o no de portal de datos abiertos; si el catálogo seguía el mismo patrón temático establecido según la Norma Técnica de Interoperabilidad (NTI); total de conjuntos de datos por CCAA; total de conjuntos de datos en salud.
- *“Conjunto de datos por formatos”*: se ha recogido en una única tabla los totales de dataset por formatos de cada CCAA; formatos reutilizables abiertos, formatos reutilizables propietarios, formatos no reutilizables (propietarios y no propietarios), formatos geolocalizables.
- *“Conjunto de dataset por temáticas”*: de cada CCAA se ha extraído el total dataset por temáticas.
- *“Materias especializadas sobre las que se está publicando en salud”*: tras la revisión de cada uno de los conjuntos de datos en salud, se generó una clasificación de materias especializadas normalizadas por la investigadora, y una subcategoría generada siguiendo la misma metodología.
- *“Formatos de datos utilizados para publicar datos en Salud”*.
- *“Calidad de los datos abiertos”*: para medir la calidad de los formatos utilizados por las CCAA para publicar en salud, se ha utilizado el modelo 5 estrellas de Berners-Lee.

Indicadores. La selección de indicadores ha sido propuesta en torno a los objetivos de esta investigación. Los indicadores han sido divididos en tres apartados principales: *portales, conjuntos de datos o dataset y calidad de los conjuntos de datos*. Los primeros de tipo cuantitativo y el tercero basado en una mezcla entre cuantitativo y cualitativo. Algunos de estos indicadores, han sido seleccionados a partir de la Norma UNE 178301:2015 (una vez revisada) y han permitido profundizar más sobre qué y cómo se

está publicando actualmente en los portales de datos abiertos de las comunidades autónomas españolas.

Limitaciones y dificultades en la recogida de datos. La principal dificultad ha sido la recogida de datos de ciertas comunidades autónomas. Ceuta y Melilla, ya que directamente no disponen de portal de datos abiertos, por lo que fueron descartadas para el objeto de estudio desde un primer momento. A medida que se recogían los datos, en algunas comunidades autónomas como Cantabria o Galicia, se ha tenido que ejecutar estrategias de búsqueda específicas adaptadas a sus buscadores, para que, a partir de los resultados obtenidos, pudiese realizarse una selección individualizada de conjuntos de datos. En el caso de Navarra, que también presentaba dificultades desde el principio, se intentó normalizar las temáticas ofrecidas en su catálogo para poder seguir la NTI, pero dado el grado de dificultad para establecer un criterio válido y adaptable, se descartó para el objeto de estudio.

2. DATOS ABIERTOS

2.1 DEFINICIÓN DE OPEN DATA O DATOS ABIERTOS

Para poder hablar de Datos Abiertos es importante entender primero el concepto datos. *¿Qué son los datos?* Los datos son la materia prima a partir de la cual se crean la información y los conocimientos. Pensemos en los datos como los lugares, las imágenes, las descripciones, las opiniones y los precios que componen la información que puede ayudarle a planificar unas vacaciones, por ejemplo. Los datos se convierten en información cuando se les proporciona un contexto. Continuando con el ejemplo anterior, los lugares, las imágenes, las descripciones, las opiniones y los precios, todos estos aspectos pueden ayudar a facilitar información relacionada con una atracción turística. Los conocimientos son lo que se obtiene de la información, y que se personalizan en función de sus necesidades. La generación de conocimientos es un proceso que consiste en transformar la información en elecciones. Basándonos en el ejemplo anterior, saber que a ningún miembro de su familia le gustan los parques de atracciones le ayudará a decidir qué lugares evitará visitar durante las vacaciones, y qué les apetecerá más. La recogida y la presentación de los datos ayuda a crear información (European Commission, s.f)

Por tanto, los datos son considerados la base para generar conocimiento nuevo en todo el mundo. La predisposición de éstos a ser reutilizados por cualquier persona, permitirá obtener numerosos beneficios a nivel mundial.



Ilustración 1. Datos, información y conocimiento

Fuente: <https://www.europeandataportal.eu/>

Según la Open Knowledge Foundation (s.f.) open data o datos abiertos son “datos que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, y que se encuentran sujetos, cuando más, al requerimiento de atribución y de compartirse de la misma manera en que aparecen”. Open Data es también una filosofía y una práctica, un movimiento social y gubernamental que se ha extendido por todo el mundo y que persigue que los datos estén disponibles de forma libre, sin restricciones o mecanismos de control (Pintos Teigeiro y Marín de la Iglesia, 2013). Para Almansa Morales (2017) “el fenómeno *open data* es una tendencia creciente construida alrededor del concepto de recopilar, ordenar y publicar los datos generados por una organización, entendiéndose un colectivo bajo las diferentes formas que puede tomar (instituciones, sociedades, empresas, Administraciones, fundaciones, etc.), de tal manera que puedan ser utilizados por una diversidad de interesados”.

Los datos abiertos, por consiguiente, son conjuntos de datos producidos o recopilados por diversos organismos que difunden para que los ciudadanos los puedan utilizar libremente de manera sencilla y cómoda. El principal objetivo de la apertura de datos es hacer accesibles de forma pública los datos, principalmente aquellos que gestiona la administración, de tal manera que cualquier persona u organización pueda manejarlos. Con este servicio, las organizaciones y las administraciones públicas aumentan su transparencia ya que el ciudadano accede a una visión real de la prestación de sus servicios. Además, la reutilización de datos abiertos por parte de empresas, entidades, asociaciones y ciudadanía en general permite la elaboración de nuevos productos y servicios que aportan valor, innovación, conocimiento y oportunidades de negocio.

Las licencias y los términos de uso de los datos abiertos están sometidas a las leyes de reutilización de la información del sector público, y en algunos casos pueden tener licencias de propiedad intelectual, aunque se tiende a abrirlas sin condiciones, siempre que se mantengan sin manipular y con la obligatoria citación de la fuente y de su última actualización.

La apertura de datos públicos comparte a su vez algunas finalidades con el acceso o la difusión de la información del sector público a través de medios electrónicos, aunque intenta superar algunos de sus límites, como los derivados de la propia actitud de las administraciones públicas frente a la difusión de la información, dando un mayor

protagonismo a los ciudadanos (Martín Delgado, 2014). También facilita la creación de información y la oferta de nuevos servicios de valor añadido basados en los datos difundidos por las administraciones públicas (Cerrillo, 2014).

No obstante, no podemos confundir transparencia con datos abiertos. La transparencia es un sistema de obligaciones a las que están sujetos los organismos públicos tras la aprobación de la Ley 19/2013, de 9 de diciembre, de Transparencia, Acceso a la Información Pública y Buen Gobierno —«*Ley de Transparencia*»— por el que hay que publicar determinada información de oficio y se debe garantizar el derecho de acceso a la información, que tanto el preámbulo de la ley como la jurisprudencia dictada hasta ahora han definido como un derecho de contenido amplio, sustantivo, subjetivo y de escasos límites. La transparencia hace referencia a la posibilidad de que la sociedad conozca y acceda a información sobre la gestión de la administración o una organización, de una forma oportuna, sin restricciones y confiable (Armstrong, 2005), mientras que los datos abiertos son eso, datos, en este caso con una licencia que permite su reutilización. Debe hacerse notar que los datos no están sujetos a la ley de propiedad intelectual, si bien la compilación de los mismos en una base de datos (por ejemplo, un dataset) si están afectados por dicha ley. La publicación de los datos, en especial de la actividad y de la gestión, es un mecanismo para que la transparencia sea verificable. Tanto mejor cuanto más relacionados entre sí y con las actividades que describen y tanto mejor cuanto más rápido se publiquen desde el momento en que se generen (FEMP, 2017).

2.2 PRINCIPIOS DE LOS DATOS ABIERTOS

En diciembre del 2007, 30 defensores del Gobierno Abierto se reunieron en Sebastopol, California, y desarrollaron un conjunto de ocho principios de datos abiertos que se plantearon como esenciales para el buen gobierno (Open Government Data Principles). Poco después, la *Sunlight Foundation* (2010) adicionó dos nuevos principios. En este contexto, los datos públicos serán considerados “abiertos” si cumplen los siguientes principios:

PRINCIPIOS DE LOS DATOS ABIERTOS	
COMPLETOS	Todo dato público debe ser puesto a disposición. Los datos públicos son datos que no están sujetos a limitaciones de privacidad, seguridad o privilegios válidos, que se rigen por otras leyes. Deben ser tan completos como sea posible.
PRIMARIOS	Los datos abiertos deben ser datos primarios originales. Se debe facilitar información sobre cómo se han obtenido los datos y dónde se encuentran los documentos originales de manera que los usuarios puedan comprobar si los datos han sido seleccionados y almacenados adecuadamente.
OPORTUNOS	Los datos deben ser puestos a disposición tan pronto como sea necesario para preservar su valor.
ACCESIBLES	Los datos deben estar tan accesibles como sea posible tanto a través de medios físicos como electrónicos para evitar los desplazamientos o la necesidad de solicitar formalmente el acceso a la información.
PROCESABLES ELECTRONICAMENTE	Los datos deben estar disponibles en formatos de uso generalizado y que faciliten su procesamiento automático.
NO DISCRIMINATORIOS	Los datos deben estar disponibles para cualquier persona, sin que sea necesario hacer una solicitud o cualquier otro trámite.
ESTANDARES ABIERTOS	Uso de estándares abiertos en la difusión de la información.
SIN LICENCIAS	El uso de los datos no debe someterse a ninguna regulación que restrinja su reutilización excepto, de forma razonable, cuando haya aspectos relativos a la privacidad o a la seguridad.
PERMANENCIA	Los datos deben estar a disposición de forma permanente. En el caso en el que se modifique la información debe indicarse.
COSTES	No se ha de establecer una contraprestación para poder acceder o reutilizar los datos y, en el caso de que se establezca, su cuantificación debe basarse en costes marginales y no en costes totales.

Tabla 1. Principios de los datos abiertos

Fuente: Cerrillo, 2017

Si bien es cierto, que estos principios de datos abiertos no se encuentran recogidos explícitamente en ninguna norma de nuestro ordenamiento jurídico, si han sido de aceptación general en el marco de los proyectos de datos abiertos desarrollados en todo el mundo.

2.3 BENEFICIOS DE LOS DATOS ABIERTOS

Con el desarrollo de la sociedad de la información, los datos abiertos son una herramienta de información muy valiosa que proporciona numerosas ventajas para diferentes sectores de la sociedad. Los principales beneficiarios son:

Beneficios para las administraciones: TRANSPARENCIA	<ul style="list-style-type: none"> • Promueve un uso inteligente y eficaz de los recursos. • Genera un gobierno transparente que aporta un mayor grado de confianza. • Facilita la interoperabilidad entre las diferentes administraciones. • Mejora la calidad de sus equipos de trabajo y servicios prestados. • Permiten detectar posibles obstáculos internos.
Beneficios para la ciudadanía: INFORMACIÓN	<ul style="list-style-type: none"> • Establece un diálogo activo, participativo y bidireccional entre gobierno y ciudadanos, principio fundamental del gobierno abierto. • Abre la posibilidad de crear nuevos servicios sociales que mejoren la vida de la ciudadanía. • Promueve una participación democrática de la ciudadanía, fortaleciéndola. • Permite que cualquier persona pueda hacer un seguimiento de la aplicación de políticas públicas y medir su eficacia.
Beneficios para las empresas: GENERAR RIQUEZA	<ul style="list-style-type: none"> • Posibilita la creación de valor económico ya que genera nuevos servicios y aplicaciones web a partir de los datos libres. • Posibilita la apertura de un nuevo mercado basado en contenidos digitales. • Posibilita la obtención de beneficios con información pública.

Tabla 2. Beneficios de la apertura de datos abiertos

Fuente: Generalitat de Catalunya. Gobierno Abierto (s.f)

Aunque los principales beneficiarios son los tres mencionados anteriormente, no son los únicos sectores que pueden hacer un uso potencial de los datos.

La apertura también puede contribuir al avance de la ciencia. La publicación de datos permite que investigadores independientes o asociados a una institución puedan tener acceso a un valioso estrato de información sobre la formulación y ejecución de políticas públicas, sobre la forma en la cual se aplican los recursos de la sociedad, sobre el impacto de los programas en el mejoramiento de la calidad de vida de las personas. Investigaciones más fundamentadas ayudan también a elevar el nivel de conocimiento que determinadas sociedades poseen sobre sí mismas y hasta sobre otras sociedades, en investigaciones internacionales, brindando condiciones para que sus actores tomen decisiones cada vez mejor justificadas.

Por otro lado, la apertura responsable y consistente de datos permite que cualquier emprendedor o grupo de emprendedores, use su creatividad para construir herramientas, soluciones y avances tecnológicos, que muchas veces escapan a los ojos de los equipos de especialistas que trabajan con esos datos dentro de la estructura gubernamental.

Y, por último, podemos encontrar a las organizaciones de la sociedad civil. La apertura de datos puede mejorar la calidad de los servicios que prestan ciertas organizaciones de cuño no gubernamental, que muchas veces son complementarias a los ofrecidos por el estado. El gobierno cuenta con una enorme cantidad de información sobre el público y servicios del interés de estas organizaciones. El acceso libre es absoluto a dichos datos les permite poner a prueba el servicio ofrecido y medir así sus resultados, mejorando la calidad de sus actividades (CEWEB, s.f)

2.4 MARCO JURÍDICO

Tanto en el ámbito español como en el europeo, se vienen desarrollando políticas relacionadas con la reutilización de información del sector público que han sido plasmadas en sus respectivos marcos legislativos. Para entender el marco legislativo de España, hay que entender que se está planteando en Europa.

2.4.1 MARCO JURÍDICO EN EUROPA

- **Libro Verde sobre la información del sector público en la sociedad de la información publicado por la UE el 20 de enero de 1999.** Un antecedente a la importante Directiva 2003/98/CE del Parlamento Europeo y del Consejo, de 17 de noviembre de 2003, relativa a la reutilización de la información del sector público, fue la publicación por la Comisión Europea, el 20 de enero de 1999, de un Libro Verde titulado “La información del sector público: un recurso clave para Europa”.

Para García Saura (2014) el objetivo del Libro Verde fue el de emprender una amplia consulta pública entre todos los operadores afectados “con miras a estudiar los principales problemas y a suscitar un debate político a escala europea”. Este libro justifica que la información del sector público es clave para

Europa y aborda cuestiones relacionadas con el acceso a la información del sector público y su comercialización. Su objetivo es tratar de incentivar a la industria europea, en clara desventaja con su homóloga estadounidense, a la hora de comercializar la información del sector público (Curto, 2017).

Según indica Martín González (2013) en este documento ya se consideraba esta información como un recurso esencial para el desarrollo económico y la creación de empleo, así como un elemento clave en el desarrollo democrático de la Unión, al facilitar la transparencia y la participación ciudadana en los asuntos públicos comunitarios. Para García García (2010) fue el detonante del debate sobre la RISP, poniendo de manifiesto que la información pública en Europa se halla fragmentada y dispersa, debido en gran parte a la existencia de diversas legislaciones nacionales sobre la manera de acceder a la información y de utilizarla. Esta situación de fragmentación en la información pública europea (al contrario que la de EE. UU.) supone una incidencia negativa grave en la competitividad de la industria europea, en particular para las pequeñas y medianas empresas (Ramos Simón et al., 2012). El tercer capítulo del Libro Verde está dedicado a cuestiones relacionadas con el acceso a la información del sector público tales como compatibilizar el acceso asequible a todos los usuarios que evite que un grupo reducido se aproveche de la información del sector público (Curto, 2017).

- **La Directiva Europea 2003/98/CE del Parlamento Europeo y del Consejo, de 17 de noviembre de 2003**, estableció un conjunto mínimo de normas para la reutilización basadas en los principios de la transparencia, en un acceso equitativo, proporcional y no discriminatorio que buscaba armonizar mínimamente las diferentes condiciones que existieran a nivel comunitario. La directiva no obligaba a los países a autorizar la reutilización, decisión que correspondía a los estados, sino que era de aplicación en los casos en que existía publicación de documentos para su reutilización, estableciendo condiciones en cuestiones relacionadas con el tratamiento de solicitudes, formatos aplicables, tarificación –aspecto que se permite, pero no se obliga con el fin de revertir el coste marginal de producción–, licencias o prohibición de los acuerdos exclusivos.

- Diez años después de la publicación de la Directiva 2003/98/CE sobre reutilización, se publica **la Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013** con el objeto de modificar la primera, tomando como referencia una consulta pública realizada sobre el conjunto de actores de la reutilización en Europa.

El documento contiene una serie de cambios con respecto a la versión anterior:

- Se introduce la autorización generalizada de la reutilización.
- Se aclara el ámbito de aplicación y se amplía a bibliotecas, archivos y museos; y, por ejemplo, se clarifica el concepto de confidencialidad comercial, en el que incluyen los «*secretos comerciales, profesionales o empresariales*».
- Se establece la obligación de que los documentos reutilizables se publiquen en formatos interoperables «mediante formatos abiertos y legibles por máquina junto con sus metadatos, con los niveles más elevados de precisión y granularidad, en un formato que garantice su interoperabilidad.
- En cuanto a las tarifas, se sigue permitiendo su cobro, aunque se indica que deberán aplicarse "*los costes marginales en que se incurra para su reproducción, puesta a disposición y difusión*". Actualmente la práctica totalidad de los datos presentes en portales de datos abiertos se ofrecen de forma totalmente gratuita.
- La normativa contempla que la reutilización se puede realizar sin sujeción a condiciones, con sujeción a condiciones basadas en licencias-tipo, o previa solicitud. Actualmente la gran mayoría de la información presente en portales de datos abiertos está sujeta a licencias Creative Commons (CC) de Reconocimiento o equivalentes. "Usa los datos para lo que quieras, citandola fuente de los mismos".

2.4.2 MARCO JURÍDICO EN ESPAÑA

- **Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público (LRISP).** La ley 30/2007, de 16 de noviembre, traspone fielmente la Directiva 2003/98/CE regulando el marco general en el cual se definen las condiciones mínimas de reutilización de la información del sector público en España. A tal efecto define reutilización como el uso de documentos que han

sido previamente elaborados y custodiados por las Administraciones Públicas en el ejercicio de sus funciones, por parte de personas físicas y jurídicas, con fines comerciales o no comerciales que no tienen que coincidir con el propósito inicial para el cual estos documentos fueron elaborados (concepto heredado de la Directiva 2003/98/CE).

Los aspectos más importantes de esta Ley son:

- Tiene como objetivo la armonización de la explotación de la información del sector público, especialmente aquella que se encuentre en soporte digital. De esta forma, de todos los documentos de libre disposición que posee el sector público, se fomentará la publicidad, así como la creación de productos y servicios de información.
 - Cada entidad pública tiene la potestad de autorizar o no la reutilización de sus documentos con fines comerciales y no comerciales. Del mismo modo, de la administración depende el establecimiento de las condiciones a las que esa reutilización estará sujeta en lo que se refiera al establecimiento de licencias de uso o a la reutilización libre de condiciones
 - Establece la posibilidad de que la AAPP solicite una contraprestación económica por la descarga/uso del documento, no obstante, se establece que el importe de esta contraprestación no podrá superar los costes totales de recogida, producción, reproducción y difusión de los documentos.
- **Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007.** Esta normativa publicada el 8 de noviembre de 2011 desarrolla la Ley 37/2007, que se describía anteriormente, sobre la reutilización de la información del sector público. En ella se establecen recomendaciones de obligado cumplimiento para que las Administraciones Públicas puedan adaptar, publicar y mantener su información reutilizable.

Sus puntos más destacables son:

- Define con precisión cuales son las obligaciones de las entidades públicas en aras de promover la reutilización. Uno de los puntos clave es que establece que todos los documentos del sector público pasan a ser reutilizables (a no ser que estén sujetos a alguna condición de exclusión), con cual la potestad de reutilizar ya no recae en una administración concreta.

- Se establece como catálogo oficial de información pública el catálogo creado en el portal datos.gov. Además, se establecen obligaciones de actualización del mismo por parte de administraciones y organismos.
- Regula las modalidades de reutilización. Se establece como normal general la puesta a disposición de la información sin condiciones específicas, pero permitiendo a los organismos establecer cuando sí sean necesarias.
- Establece la regulación del régimen aplicable a documentos reutilizables sujetos a derechos de propiedad intelectual, industrial o que contengan datos personales.

Es importante resaltar de este decreto, que también indica que en el momento que se pone a disposición de los reutilizadores para su uso, se están cediendo los derechos necesarios para desarrollar cualquier actividad de reutilización autorizada de forma gratuita, y no exclusiva.

- **Ley 18/2015, de 9 de julio, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público.** Es la incorporación al ordenamiento jurídico español de los cambios que ha introducido la Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013, en el régimen de reutilización de documentos del sector público (detallada anteriormente).

Otras normas que inciden en el intercambio y reutilización de información son:

- **Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad** en el ámbito de la Administración Electrónica, que prevé la interoperabilidad como capacidad de los sistemas de información para posibilitar el intercambio de la misma.
- **Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno**, que recoge entre los principios generales de la publicidad activa y del Portal de Transparencia, el de la reutilización de la información pública.
- **Norma UNE 178301 Ciudades Inteligentes. Datos Abiertos (Open Data)** es la primera norma que establece un conjunto de requisitos, para la reutilización de Datos Abiertos u Open Data elaborados o custodiados por el sector público, en

el ámbito de las Ciudades Inteligentes. Los datos abiertos establecen mecanismos de transparencia y mejoran la interoperabilidad entre la Administración y entidades colaboradoras, aumentando así la rapidez, el acceso a la información y, por tanto, la inteligencia de las ciudades. Además, facilita la reutilización de los datos elaborados o custodiados por el sector público (open data) a los beneficiarios de esa información, como son las empresas proveedoras de servicios públicos o los ciudadanos, teniendo en consideración la aplicación de la ley de protección de datos de carácter personal. Además, sirve de referencia a la Administración en la implantación y gestión de proyectos de datos abiertos. Por otro lado, evalúa la madurez del proyecto de datos abiertos de la ciudad inteligente, gracias a indicadores y métricas relativos a sostenibilidad, calidad, eficacia y eficiencia del proyecto.

Hay que puntualizar, de que, aunque existe más legislación vigente, en este trabajo solo se han especificado aquellas normas que se han considerado de mayor relevancia para comprender el objeto de estudio de este trabajo. La cronología completa de la evolución del Linked Open Data que estará disponible en el ANEXO I de este trabajo.

2.5 CONDICIONES DE USO Y LICENCIAS

El desarrollo de un proyecto de datos abiertos debe tener en cuenta los requisitos establecidos en un amplio conjunto de normas estatales, autonómicas y, en su caso, locales que regulan la administración electrónica, la transparencia, el acceso y la reutilización de la información pública. Estas normas deben considerarse a la hora de diseñar cualquier proyecto de datos abiertos, pero también a la hora de poner estos datos a disposición de los ciudadanos (Cerrillo, 2014).

Ante esta la pluralidad normativa cada portal de datos abiertos debería concretar las condiciones específicas a las que se somete la reutilización¹⁰ de los datos abiertos que se difunden a través de dicho portal mediante la publicación de avisos legales¹¹, el

¹⁰ Reutilización de datos “uso de los datos que se encuentran en poder de las administraciones públicas, con fines comerciales o no comerciales, siempre que dicho uso no constituya una actividad administrativa pública” (Artículo 3 de la LRISP)

¹¹ Aviso Legal “documento mediante el cual las administraciones públicas informan a los ciudadanos de las normas que regulan la difusión y reutilización de los datos y los efectos jurídicos de dichas actividades” (Cerrillo, 2014)

establecimiento de licencias para la reutilización, la inclusión de licencias de uso de los portales o la previsión de códigos de conducta (Cerrillo, 2014).

También es necesario que los ciudadanos, cuando utilizan el portal de datos abiertos para acceder a datos que posteriormente reutilizarán, tengan seguridad jurídica y también que sus derechos se vean plenamente garantizados y sus obligaciones claramente definidas (Cerrillo, 2014). Y, asimismo las administraciones públicas deben velar por salvaguardar sus derechos y bienes y los de terceros (por ejemplo, en relación con los datos personales o con la propiedad intelectual).

En España cada portal de datos abiertos determina un tipo de licencia que acompañe a la publicación de los conjuntos de datos, pudiendo otorgar libertad para compartir, distribuir, adaptar o utilizar los datos, así como para crear obras derivadas. No obstante, existen diferentes licencias dependiendo de las condiciones de reutilización impuestas a dichos datos abiertos (datos.gob.es, s.f)

Según los fines comerciales que se desee otorgar a la información reutilizada, se distinguen:

- *Licencias de uso comercial:* permiten un aprovechamiento económico de las obras derivadas.
- *Licencias de uso no comercial:* no permiten un fin comercial tanto de los datos originales como de las posibles obras derivadas de ellos.

A su vez, también pueden clasificarse según el acceso que se brinda a los datos:

- *Licencias abiertas:* apenas impone restricciones y permiten difundir, reutilizar y publicar los datos, con fines comerciales o no comerciales. Tan solo se podría requerir que el reutilizador reconozca la fuente de los datos originales y publique la obra nueva bajo la misma licencia que la información de origen. Estas son las licencias que más fomentan los principios de Open Data.
- *Licencias con condiciones específicas:* aquellas licencias que se establecen en un organismo deben ser uniformes en su tipología y, además, deben incorporar al menos, la información relativa a la finalidad concreta de la licencia, indicando si su uso va a ser comercial o no, los motivos para los que se concede la

reutilización, así como la duración de la misma y las obligaciones de todas las partes. También las responsabilidades de uso, y si es de carácter gratuito o no; en caso de que no, indicará la tasa o precio público aplicable (PINTOS, 2012).

- *Licencia previa solicitud para la reutilización de los datos:* normalmente se aplica a todos los documentos que no se encuentren en formato digital. Y es obligatorio que los organismos públicos tengan disponible en su página web la información referente a las tasas o precios públicos aplicables para los documentos reutilizables bajo esta modalidad (PINTOS, 2012).

En la búsqueda de una propiedad intelectual más flexible, en 2002 nació la corporación sin ánimo de lucro Creative Commons (CC). Esta entidad creó una serie de licencias que permiten compartir los trabajos bajo unos términos flexibles normalizados. De las seis licencias disponibles, cuatro de ellas son aplicables en open data (datos.gob.es. s.f):

- **Creative Commons Zero (CC0):** ningún tipo de restricción, renunciando el creador a cualquier derecho sobre la obra, según permita la ley.
- **Creative Commons Reconocimiento (CC BY):** solo requiere que el reutilizador haga referencia a la fuente o autor original.
- **Creative Commons Reconocimiento-NoDerivadas (CC BY-ND):** permite la reutilización y difusión siempre y cuando se haga referencia al autor de la misma y la obra no sufra cambio o alteración alguna.
- **Creative Commons Reconocimiento-NoComercial (CC BY-NC):** es posible alterar o difundir la obra original siempre y cuando se haga referencia al autor de la misma, careciendo de fines comerciales. La obra derivada no está obligada a mantener la misma licencia que la obra original:

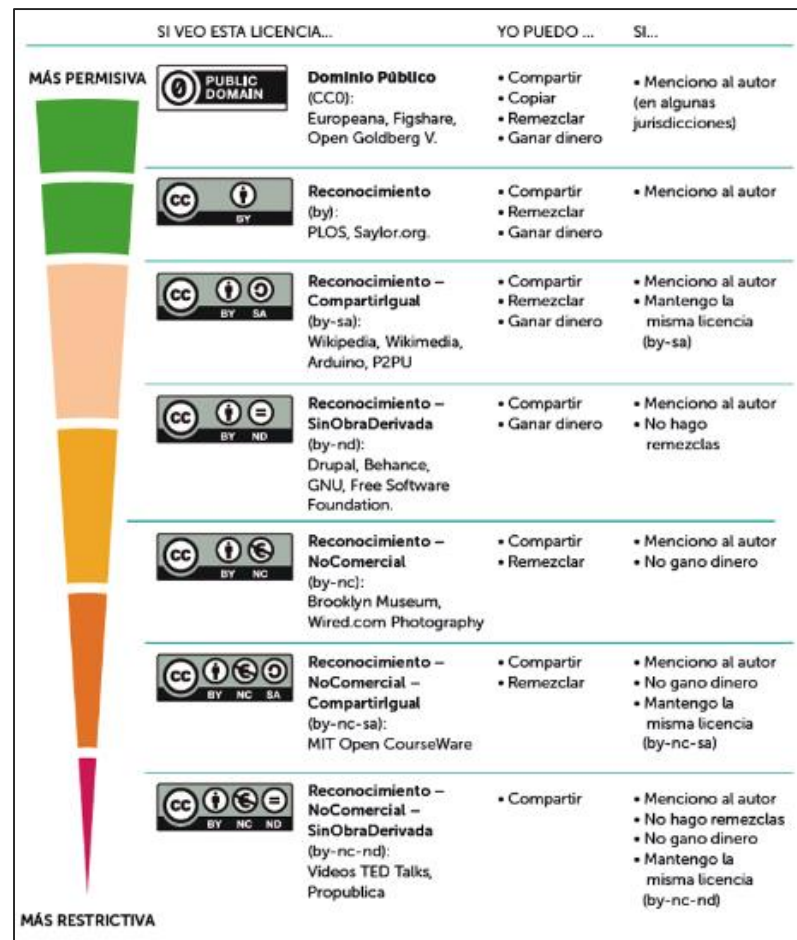


Ilustración 2. Clasificación de licencias Creative Commons

Fuente: Atresmedia, 2015

Hasta la aparición de la versión 4.0, estas licencias se centraban principalmente en los contenidos y no en los datos. Por este motivo, en 2009, la organización OKFN (Open Knowledge Foundation Network) desarrolló las primeras licencias específicas para datos abiertos, las licencias Open Data Commons:

- **Open Data Commons Public Domain Dedication and License (PDDL)**: permite difundir, reutilizar o adaptar los datos sin restricción alguna.
- **Open Data Commons Attribution License**: se exige la referencia a la autoría o fuente de los datos para la reutilización de la información.
- **Open Data Commons Open Database License (ODbL)**: permite la reutilización de los datos siempre que se reconozca la autoría de la información original; se mantenga la misma licencia en las obras derivadas las cuales puede restringir su uso si, además, se distribuye una versión sin dichas restricciones de uso.

En la Administración General de Estado se ha realizado un esfuerzo para homogeneizar las condiciones de reutilización de la información pública mediante el Real Decreto 1495/2011. Este Real Decreto establece que la información disponible es reutilizable por defecto y establece de manera precisa las condiciones generales básicas que se aplican en la reutilización de la información pública. Estas condiciones se pueden resumir de la siguiente forma:

1. No desnaturalizar el sentido de la información.
2. Citar la fuente de los documentos.
3. Mencionar la fecha de la última actualización, siempre que esté incluida en el documento original (es responsabilidad del organismo el mantener la información actualizada).
4. Conservar y no alterar ni suprimir los metadatos relativos a la fecha de actualización y a las condiciones aplicables a la dicha reutilización.

Dichas condiciones estarán accesibles por medios electrónicos y vincularán a cualquier agente reutilizador por el mero hecho de hacer uso de los documentos. En el caso de que una Administración pública quiera establecer otras condiciones, la LRISP prevé que estas deben ser claras, justas y transparentes, no deben restringir las posibilidades de reutilización ni deben limitar la competencia y no deben ser discriminatorias para categorías comparables de reutilización (Cerrillo, 2014).

La LRISP también prevé la posibilidad de establecer un precio o contraprestación económica por la reutilización. No obstante, a fin de facilitar al máximo la reutilización es recomendable no establecer ninguna, tal como sucede en la mayoría de los portales de datos abiertos (Cerrillo, 2014).

2.6 CALIDAD DE LOS DATOS ABIERTOS

Según la Declaración sobre gobierno abierto adoptada por la iniciativa *Open Government Partnership*, los cincuenta gobiernos firmantes –entre los que se encuentra España– se comprometen a *“proporcionar activamente información de alto valor,*

incluidos los datos primarios, de manera oportuna, en formatos que el público pueda encontrar, comprender y utilizar fácilmente, y en formatos que faciliten su reutilización.”

Hoy en día disponemos cada vez mas de fuentes de datos a nuestro alcance. Sin embargo, paradójicamente, esto no quiere decir que los datos sean más fáciles de usar, simplemente que existe más cantidad de datos publicados. El éxito de los sistemas de datos abiertos depende, más que del simple acceso, de la valoración de la calidad de los mismos (Jannssen, et.al., 2012). Los autores Ferrer-Sapena, Peset y Aleixandre Benavent (2011) puntualizan que la puesta a disposición de los datos no tiene sentido si posteriormente no se trata para generar conocimiento.

La mera existencia de datos abiertos no implica su reutilización. Maseda-Seco, Bueno-de-la-Fuente y Méndez (2017) apuntan que “no toda la información pública disponible o publicada en la Web es información abierta válida para su reutilización”. Además de publicar los datos, hay que dar acceso a los mismos y dotarlos de una estructura que permita su reutilización (Garriga-Portolà, 2011). Gértrudix, Gertrudis-Casado y Álvarez-García (2016) explican que hay cierta desconfianza de los ciudadanos respecto de las acciones que realizan las instituciones públicas para acceder, utilizar y reutilizar datos abiertos. Maseda-Seco, Bueno-de-la-Fuente y Méndez (2017) consideran que la clasificación de Berners-Lee (2006) es una herramienta útil para analizar la apertura y potencial de reutilización, explotación y consumo de los datos.

2.6.1 MODELO DE CALIDAD: LAS CINCO ESTRELLAS DE BERNERS-LEE

Tim Berners-Lee es el inventor de la Web e iniciador de los Datos Enlazados (Linked Data). En mayo de 2010 durante el evento Gov 2.0 Expo 2010¹² expone el llamado “*Modelo de Calidad de 5 Estrellas*”. Este modelo o esquema establece cinco niveles de calidad etiquetados con estrellas según la facilidad de uso por parte de los consumidores de datos:

¹² Para más información se puede consultar: <https://www.youtube.com/watch?v=ga1aSJXCFe0>

ESTRELLAS	SIGNIFICADO Y EJEMPLOS
★	La información está en la web (en cualquier formato no estructurado). Esto implica que, para utilizar los datos, se requiere la aplicación de técnicas o esfuerzo extra. <i>Ejemplo: formatos de imagen, vídeo o música. También los PDF.</i>
★★	Los datos están publicados en formatos estructurados (legibles por máquinas) lo que facilita su tratamiento; no obstante, están en formato propietario lo que puede exigir la compra de una licencia para poder trabajar con la información. <i>Ejemplo: XLS (Microsoft Excel).</i>
★★★	Los datos están publicados en la web en formato estructurado y, además, son abiertos y no propietarios. Cualquier persona puede acceder y utilizar los datos de forma sencilla. <i>Ejemplo: CSV.</i>
★★★★	Utilización de URI (Identificador de Recursos Uniforme), que permite que los usuarios pueden reutilizar partes de los datos y que estos puedan vincular con los de cualquier otro. Los datos ya están "en la web" y van a adquirir significado dependiendo de las etiquetas que se utilicen para la confección del documento.
★★★★★	Nivel linked data: Los datos originales enlazan, utilizando URI, con otros datos nuevos de otros suministradores. Esta vinculación de datos permite dotar a los datos originales de un nuevo contexto.

Tabla 3. Esquema del Modelo 5 Estrellas Berners-Lee

Fuente: Berners-Lee (2010)

Con este esquema, Tim Berners-Lee muestra que la apertura de datos tiene niveles, ofreciendo un diagrama de clasificación que permite apreciar las características de cada nivel de apertura, siendo el nivel óptimo es de las cinco estrellas. Este esquema de cinco estrellas muestra la calidad a efectos de reutilización de los formatos en los que se encuentra codificada la información (Curto, 2017). Es evidente que cuanto mejor estructurados y enriquecidos estén los datos más fáciles serán su reutilización y la construcción de aplicaciones para su tratamiento automatizado.

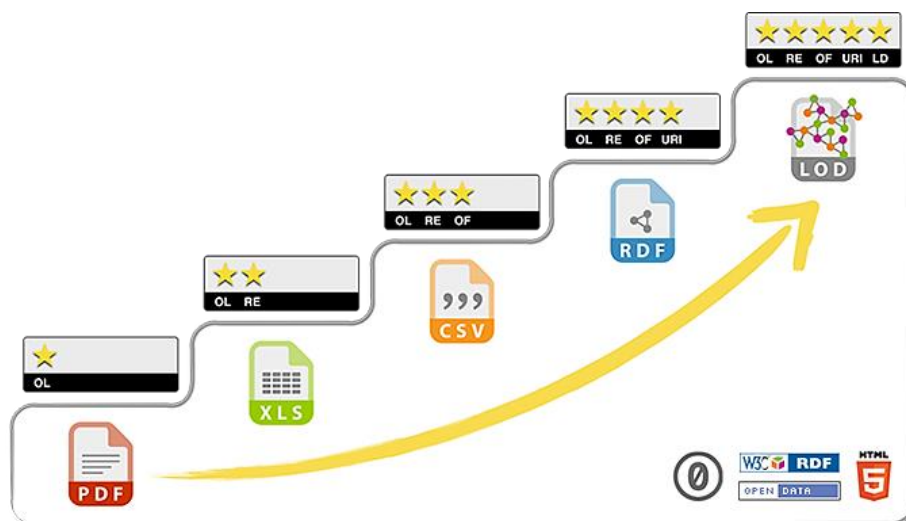


Ilustración 3. Esquema 5 estrellas Tim Berners-Lee

Fuente: <https://5stardata.info/es/>

2.6.2 NORMA 178301:2015

La Asociación Española de Normalización y Certificación (AENOR) publicó en el año 2015 la primera norma para las *smart cities*, se trata de la UNE 178301 sobre Datos Abiertos (Open Data), donde se establecen los requisitos para la reutilización de datos elaborados o custodiados por el sector público. Con la publicación de esta norma, se pretenden que los ciudadanos tengan una imagen clara y cuantificada de la apertura de los datos que gestionan las Administraciones y organismos del sector público que gobiernan en las ciudades. Al tratarse de una norma técnica, hay que tener en cuenta que es un documento elaborado desde el consenso, teniendo en cuenta tanto la experiencia como el desarrollo tecnológico, además de ser aprobado por un organismo de normalización reconocido.

La UNE 178301:2015 nos indica cómo evaluar la madurez de un proyecto de datos abiertos a través de una serie de indicadores y métricas relacionados con criterios de sostenibilidad, calidad, eficacia y eficiencia de la misma iniciativa.

Las métricas se organizan en forma de dominios y dimensiones para facilitar su estructuración y comprensión:

- *Dominio estratégico*: establece los criterios para evaluar la capacidad y ejecución de una política de apertura de datos. Sus dimensiones son la estrategia, el liderazgo, el compromiso de servicio y la sostenibilidad económica.
- *Dominio legal*: establece los criterios para evaluar la existencia y verificación de leyes y normativas que facilitan la ejecución de las políticas de datos abiertos. Su dimensión es la legal, que se subdivide en normas externas e internas y las condiciones de uso y licenciamiento.
- *Dominio organizativo*: establece los criterios para evaluar la capacidad para ejercer adecuadamente actividades que permitan desarrollar de una estrategia de apertura de datos. Las dimensiones son la organizativa (ver si existe una unidad responsable de la apertura de datos, equipo de trabajo y su capacitación, inventario y prioridad) y la medición, ésta subdividida en medición de cumplimiento de proceso y medición del uso e impacto

- **Dominio técnico:** establece los criterios para evaluar aquellas actividades que garanticen la disponibilidad de los datos en cada momento. Sus dimensiones son la disponibilidad (presencia en el catálogo de información pública, conjuntos de datos documentados, categorización y búsqueda, disponibilidad y URLs persistentes y amigables), el acceso (accesibilidad / no discriminación, gratuidad y sistemas de acceso), la calidad de los datos (datos primarios, datos completos, datos documentados, datos técnicamente correctos, datos georreferenciados y datos enlazados) y la actualización (proceso de actualización, frecuencia de actualización y ampliación de conjuntos de datos).
- **Dominio económico y social:** establece criterios para evaluar los mecanismos que relacionan a los productores de los datos con los reutilizadores, la compartición de estructuras comunes que fomenten la aplicación de los datos en la producción de nuevos bienes y servicios, el grado de implicación del organismo en el estímulo y ayuda a la labor de los agentes reutilizadores, el grado de escucha y adaptación a las demandas y el nivel de diálogo establecido. Las dimensiones son la reutilización (cantidad de datos, formato de los datos, vocabularios de datos estándares), la dimensión participación y colaboración (transparencia, participación y colaboración, resolución de quejas y conflictos, fomento de la reutilización e Iniciativas de reutilización desarrolladas).



Ilustración 4. Representación de los dominios de la Norma UNE 178301:2015 sobre Datos Abiertos

Fuente: elaboración propia

2.7 FORMATOS DE DATOS ABIERTOS

Uno de los puntos clave de toda política de datos abiertos es que estos sean presentados en un formato que permita su reutilización por parte de terceros. Existen diferentes formatos para fomentar la reutilización de datos abiertos. La elección para la reutilización de un formato u otro dependerá del usuario final, que decidirá cuál es el formato que más se adecua al uso que quiera dar a dichos datos

Por ello, una recomendación habitual es el empleo de formatos estándar abiertos que permitirá que los datos sean tratados con diversos tipos de herramientas sin necesidades de tener que adquirir licencias de uso.

Existen muchos formatos de ficheros, pero no todos son reutilizables, ni abiertos. Los *formatos reutilizables* son formatos de datos estructurados, que pueden ser leídos o interpretados por máquina. Si además estos formatos no son propietarios se consideran formatos abiertos. Se entienden por *abiertos* los formatos en los que su estructura es conocida de una forma pública y la estructura/especificación suele ser aprobada por organizaciones que trabajan con estándares. Mientras que los formatos *propietarios*, son formatos que requieren herramientas que no son públicas.

Por el contrario, los formatos *no reutilizables* no pueden ser procesados automáticamente, pero la publicación en estos formatos suele darse con bastante frecuencia.

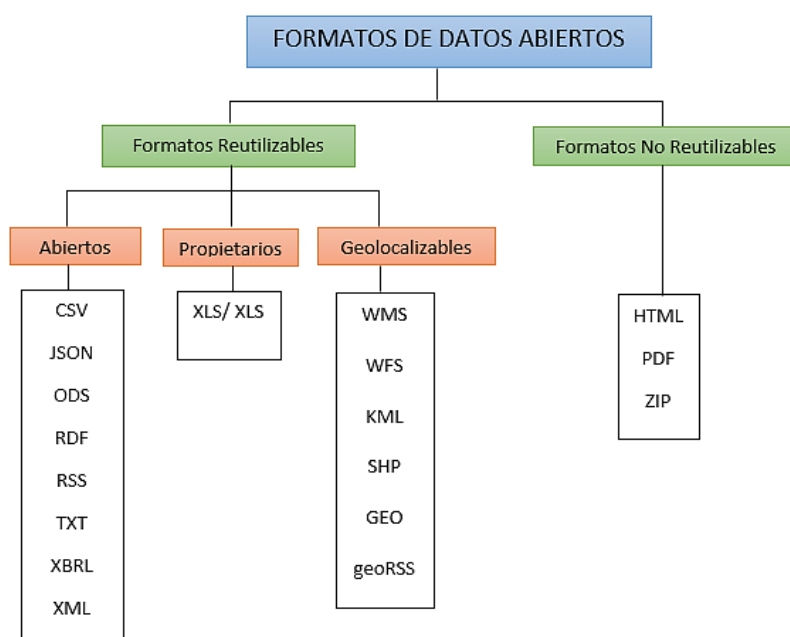


Ilustración 5. Esquema de formatos de Open Data

Fuente: elaboración propia

A continuación, se detallan cada uno de ellos:

FORMATOS REUTILIZABLES	
ABIERTOS	
CSV	Valores separados por coma. Los ficheros CSV son un tipo de documento en formato abierto sencillo para representar datos en formato de tabla. Las columnas se separan por comas (o punto y coma) y las filas por saltos de línea.
XML	Este formato tiene una importante función en el proceso de intercambio, estructuración y envío de datos en la Red. Describe los datos de tal manera que es posible estructurarlos utilizando para ello etiquetas que no están predefinidas. De esta manera los datos están delimitados, lo que favorece la interoperabilidad de los mismos.
ODS	Formato de datos abierto y estándar, se emplea para el almacenamiento de hojas de cálculo que muestran información en celdas organizadas en filas y columnas. Cada una de esas celdas contiene datos o fórmulas con referencias - relativas o absolutas - a otras celdas.
JSON	Es un formato de texto ligero para el intercambio de datos. JSON es un subconjunto de la notación literal de objetos de JavaScript aunque hoy, debido a su amplia adopción como alternativa a XML, se considera un formato de lenguaje independiente.
RSS	Es un formato XML para la distribución de contenidos de páginas web. Facilita la publicación de información actualizada a los usuarios suscritos a la fuente RSS sin necesidad de usar un navegador, utilizando un software especializado en este formato.
TXT	Documento de texto plano. Puede utilizar diferentes codificaciones de caracteres.
RDF	Es un modelo universal basado en XML que permite intercambiar y enlazar a través de diferentes aplicaciones datos y recursos sin que pierdan su significado, lo que facilita la reutilización y enriquecimiento de los recursos en la Web
XBRL	Se creó para simplificar la automatización del intercambio de información financiera mediante el uso del lenguaje XML.
PROPIETARIOS	
XLS/XLSX	Son un formato propietario de Microsoft que muestran la información en celdas organizadas en filas y columnas, y cada celda contiene datos o fórmulas, con referencias relativas o absolutas a otras celdas.
GEOLOCALIZABLES	
WMS SHP WFS GEO KML geoRSS	Estos formatos se usan para publicar información geográfica de forma que se pueda reutilizar. Los datos son publicados de forma georeferenciada y muchos de ellos son los formatos/servicios a utilizar para cumplir con la normativa europea (INSPIRE).

Tabla 4. Formatos Reutilizables utilizados en la apertura de datos.

Fuente: Portal de datos abiertos Ayuntamiento de Madrid

FORMATOS NO REUTILIZABLES	
HTML	Se trata de un lenguaje que nos sirve para estructurar textos y presentarlos de forma que puedan ser visualizados a través de navegadores web. Se trata de uno de los formatos más extendidos y populares de información accesible a través de un navegador. Ofrece para el reutilizador la ventaja de establecer una estructura de documento que permite su tratamiento de forma automática (Pintos, 2012)
PDF	Es un formato de almacenamiento de documentos digitales independiente de plataformas de software o hardware. Este formato es de tipo compuesto que incluye imágenes vectoriales, mapa de bits y texto.
ZIP	Es un formato de compresión sin pérdida, muy utilizado para la compresión de documentos, imágenes o programas.

Tabla 5. *Formatos no reutilizables en materia de open data*

Fuente: Portal de datos abiertos Ayuntamiento de Madrid

Según el modelo de 5 estrellas de Tim Berners-Lee los formatos utilizados en Open Data se clasificarían de la siguiente manera:


CLASIFICACIÓN DE FORMATOS SEGÚN EL MODELO 5 ESTRELLAS TIM BERNERS-LEE	
	Djvu, DOC, DXF, GZIP, IMG, JPEG, ODP, ODT, PDF, PNG, PostScript, PPT, RTF, Sold, TIFF, Visio, ZIP
	DBF, GDB, DGN, DWG, ECW, ELP, GeoPDF, MDB, MPP, Pc-Axis, PGP, RASTER, SHP, XLS
	API, ASCII, ATOM, Calendar, CSV, CSW, DOCX, ePub, GeoJSON, GeoRSS, GML, GPX, HTML, JSON, KML, KMZ, LAS, MP4, NetCDF, OCTET-STREAM, ODS, RSS, SCORM, SMIL, SOAP, SubRip, TBX, Texto plano, TMX, TSV, vCard-texto, vCard-xml, Voice-XML, WMS, WMS-XML, WFS, WCS, XBRL, XHTML, XLSX, XML, XML-APP, Z39.50
	JSON-LD, N3, RDF-XML, RDF-N3, RDF-Turtle, SPARQL, SPARQL-XML, SPARQL-JSON, TURTLE

Tabla 6. *Clasificación de formatos según el modelo 5 Estrellas de Tim Berners-Lee*

Fuente: datos.gob.es, s.f

3. DATOS ABIERTOS EN SALUD

3.1 PANORAMA ACTUAL

El ámbito de la salud ha estado hasta hace no mucho tiempo ajeno a la apertura de los datos, entre otros motivos, por los riesgos derivados del tratamiento de información muy sensible. Pero la apertura de datos es ya un hecho y el sector salud es uno de los ámbitos en el que esta apertura puede tener un mayor impacto en los próximos años.

Actualmente estamos asistiendo a una auténtica revolución en el ámbito de la salud. Este hecho se puede comprobar con la aparición cada vez mayor de noticias (además de las declaraciones al respecto o en informes oficiales) sobre cuestiones de salud, evoluciones tecnológicas y tratamiento de datos. Los cambios culturales, demográficos, sociales (envejecimiento de la población, aumento de la cronicidad y comorbilidad, etc.) y tecnológicos, tienen un reflejo evidente en esta realidad. Con la apertura de datos sanitarios se pretende contribuir a posibilitar lo que se consideran los principales retos de los sistemas sanitarios en los próximos años (Rovira, 2016):

- ✓ Reforzar la prevención y la salud pública: se busca una medicina más proactiva y no tanto reactiva. Se pretende retrasar la aparición de las enfermedades o de las situaciones de dependencia. Y ello implica una mayor “responsabilización” del ciudadano en la gestión de su propia salud.
- ✓ Asegurar la continuidad asistencial a lo largo de la vida: con la finalidad de conseguir una atención sanitaria continuada y de calidad. Esto supone, entre otros, una mayor disponibilidad de la información sobre salud de las personas.
- ✓ Resaltar la necesidad de conseguir una mayor coordinación socio-sanitaria: se trataría de superar el actual marco de la sanidad para establecer puentes con las estructuras sociales (lo que implica también un mayor trasvase de datos).

Hay que tener en cuenta que el sector salud genera un volumen ingente de datos y por muchos actores, y que la cantidad de datos relacionados con la salud y la sanidad ha crecido de forma exponencial en los últimos años. En España, como ya se ha mencionado anteriormente, el volumen mayor de datos se concentra en manos de las Administraciones públicas (principales prestadores de la asistencia sanitaria; con

información asistencial, datos administrativos, pero también de costes y otros). Aunque cada vez generan mayor información otros proveedores de salud, la industria sanitaria (en particular, la industria farmacéutica), etc. Por otra parte, también va creciendo el número de datos de salud que generan los dispositivos móviles y las aplicaciones sobre salud y bienestar, junto con las wearables.

A nivel europeo el país más avanzado en esta materia es Reino Unido, con el portal de datos del Servicio público de salud (NHS)¹³. No solo presenta datos e indicadores estadísticos, sino también valoraciones de cada centro sanitario, coste de tratamiento, resultados de tratamientos clínicos, incidencia de enfermedades, etc. Con ello, los ciudadanos pueden ver el impacto de sus impuestos en sanidad y comparar el nivel de asistencia médica de cada área, al igual que dota a las Administraciones de información para la toma de decisiones y mejora de los servicios.

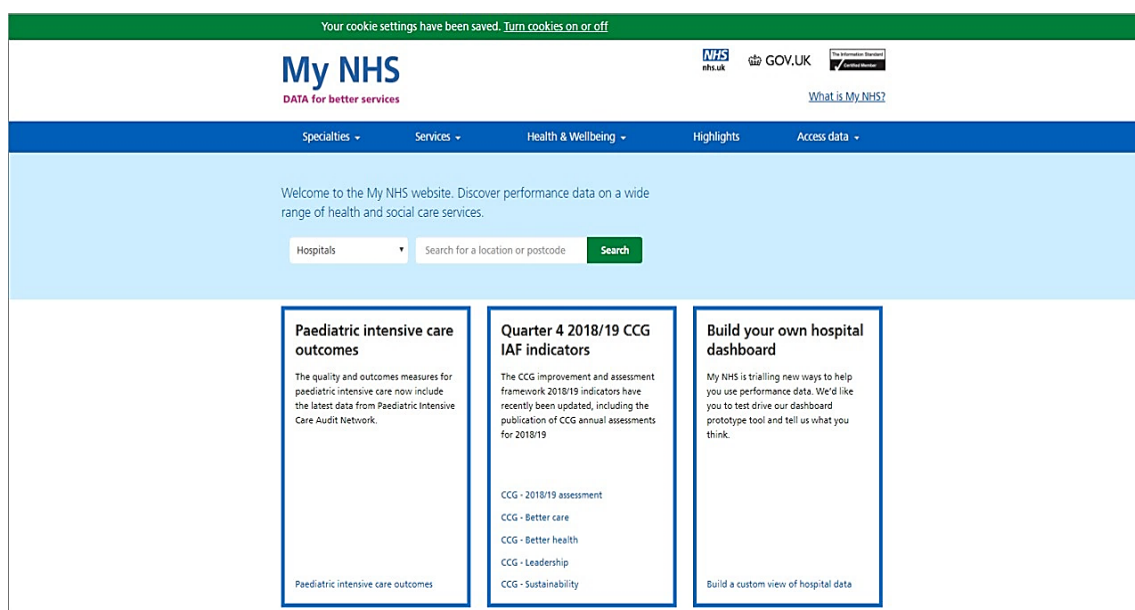


Ilustración 6. Pantalla principal del Portal de datos abiertos de la NHS de Reino Unido

Fuente: <https://www.nhs.uk/service-search/performance/search>

¹³ Para más información: <https://www.nhs.uk/service-search/performance/search>

3.2 PANORAMA NORMATIVO PARA EL ACCESO Y TRATAMIENTO DE DATOS EN SALUD

En España, aunque el panorama legislativo en materia de tratamiento de datos de salud es disperso y se encuentra fragmentado¹⁴, podemos considerar que las dos normas básicas de referencia serían la nueva Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos de Carácter Personal (LOPD) y la Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica, que regula entre otros el acceso a la historia clínica. Pues bien, ambas normas están pensando en un tratamiento de los datos de salud en un contexto de asistencia sanitaria que podemos denominar “clásico” y, por tanto, en un tratamiento que se realiza con la finalidad básicamente de atención sanitaria en una relación individual médico-paciente (con algunas excepciones, como la utilización por terceros con funciones de evaluación, acreditación o planificación del sistema; o con fines epidemiológicos, de salud pública y de investigación o docencia, fundamentalmente con datos disociados). En cualquier caso, esta regulación está basada en un uso de los datos de salud en el contexto sanitario previo a la introducción masiva de las tecnologías de información y comunicación (TIC) en este ámbito, a la época big data o a la cultura del open data y, por lo tanto, no dan una respuesta satisfactoria a los interrogantes que se pueden plantear en el actual escenario. Ambas normas deben ser objeto de una revisión importante.

Por su parte, la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno (LTAIPBG), en sus artículos 5.3, 15 y la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público (LRISP), en sus artículos 3.3.a y 3.4, contienen las excepciones clásicas al acceso y la reutilización de datos en poder de las Administraciones públicas cuando se trate de datos de carácter personal y, si son sensibles (como es el caso que nos ocupa), se inclinan por la disociación de datos para permitir el acceso y reutilización de los mismos.

¹⁴ Para un conocimiento más en detalle de la normativa española en el tratamiento de datos de salud, así como en transparencia, acceso y reutilización de la información sanitaria puede consultarse Valero Torrijos (2014) y Aberasturi Gorriño (2013).

3.3 BENEFICIOS

El open data debería permitir a los gobiernos una mejora en la toma de decisiones en el sector salud (aumentando la eficiencia del sistema y reduciendo costes, entre otros) y un incremento de la transparencia en la gestión pública. Por lo que hace al paciente, al estar mejor informado, involucrarse más en la toma de decisiones. Por otra parte, se dice que los pacientes más informados y formados deberían responder mejor a los tratamientos, lo que repercutiría en una mejora del sistema sanitario. En definitiva, una mayor participación, responsabilidad y compromiso por parte del paciente (“empoderamiento” del paciente). Por último, también se destaca la creación de nuevos modelos de negocio con base en el open data. En el ámbito sanitario permitiría crear nuevas herramientas y aplicaciones de salud basadas en los datos abiertos (ej., para evaluar la calidad de los servicios de salud, costes de prestaciones, etc.).

No obstante, a pesar de las grandes ventajas que supone la reutilización del open data sanitario, paralelamente también conlleva nuevos desafíos; en especial, el almacenamiento, tratamiento en tiempo real e integración del gran volumen de dataset existente. Si bien es cierto que, gracias a los avances tecnológicos y los servicios en la nube, cada vez es más fácil sortear dichos obstáculos. Al fin y al cabo, la información ya no se recopila en un lugar único y esto permite que diferentes agentes tengan acceso a ella, agilizando su reutilización.

3.4 OBSTÁCULOS EN LA IMPLANTACIÓN

Los obstáculos con los que se encuentra el open data también han sido señalados (entre otros, por el NHS inglés)

- Las barreras institucionales, burocráticas y culturales a la apertura de datos. La cultura de los datos abiertos requiere un cambio de percepción e incluso cambio cultural en algunos casos. Se está pidiendo a las administraciones, proveedores de salud y pacientes que se adapten a una nueva forma de recolectar y usar los datos.
- Protección de la privacidad. En el ámbito de la salud se tratan datos sensibles y, existe cierta desconfianza hacia un uso incorrecto de estos datos que puede ser perjudicial para la persona.

- También existe desconfianza a que la apertura de datos lleve realmente a una toma de decisiones adecuada (la necesidad de un “buen análisis”). Buenos datos no siempre significan buenas decisiones, pues al final siempre hay un componente de análisis e interpretación humana.
- Y, además se han puesto de relieve los obstáculos relacionados con estándares técnicos y la interoperabilidad, que puede producir que la apertura de datos no sea efectiva. Se habla en este sentido de la necesaria “liquidez de los datos”, esto es, de la capacidad de los datos de fluir fácilmente entre proveedores, pacientes y otros grupos que intervienen en la asistencia sanitaria.

Un reflejo de lo anterior son las dificultades con las que se encuentra en la práctica la puesta en marcha de proyectos de open data en el ámbito sanitario¹⁵. En España (en concreto, en Cataluña) ha habido una experiencia incipiente de apertura de datos, sobre todo de reutilización de datos de salud del sistema público catalán, que se denominó en su origen Visc+ y que fue objeto de una gran controversia a nivel político y social¹⁶. El proyecto se fue modificando con el tiempo y “restringiendo” antes las críticas de una posible “venta” de datos de las historias clínicas públicas¹⁷ y en la actualidad se encuentra como tal paralizado y ha sido sustituido por el Programa público de analítica de datos para la investigación y la innovación en salud (PADRIS¹⁸) de la Agencia de Calidad y Evaluación de Datos de Cataluña (AQuAS) PADRIS tiene la misión de poner a disposición de la comunidad científica datos sanitarios para impulsar la investigación, la innovación y la evaluación en salud, facilitando su acceso y reutilización de acuerdo con el marco legal y normativo, los principios éticos y de transparencia (Agència de Qualitat i Avaluació Sanitàries de Catalunya, s.f.).

¹⁵ La dificultad de tratar con los datos de los historiales médicos por las suspicacias que suscita se puso de relieve también en el Reino Unido a propósito del programa Care.data en el año 2013 por parte del NHS, que buscaba centralizar datos médicos y sociales de los pacientes y le permitía compartir información con terceros fuera del NHS. A instancia de la National Data Guardian for Health and Care (que puso en cuestión el modelo de consentimiento y el opt-out propuesto), el NSH decidió cerrar el programa y trasladar la labor iniciada a la National Information Board para reforzar la confianza ciudadana en esta materia.

Véase <https://www.england.nhs.uk/ourwork/tsd/care-data/> y <https://www.england.nhs.uk/digitaltechnology/info-revolution/>

¹⁶ Se trataba de una licitación de un contrato de colaboración público-privada para la implantación y operacionalización de un modelo de gestión para dar valor a la información del sistema sanitario catalán en el marco de políticas públicas. Ya la propia denominación (oscura) fue criticada.

¹⁷ Un análisis crítico del proyecto puede consultarse en Serrano Pérez (2015).

¹⁸ Programa público de analítica de datos para la investigación y la innovación en salud.

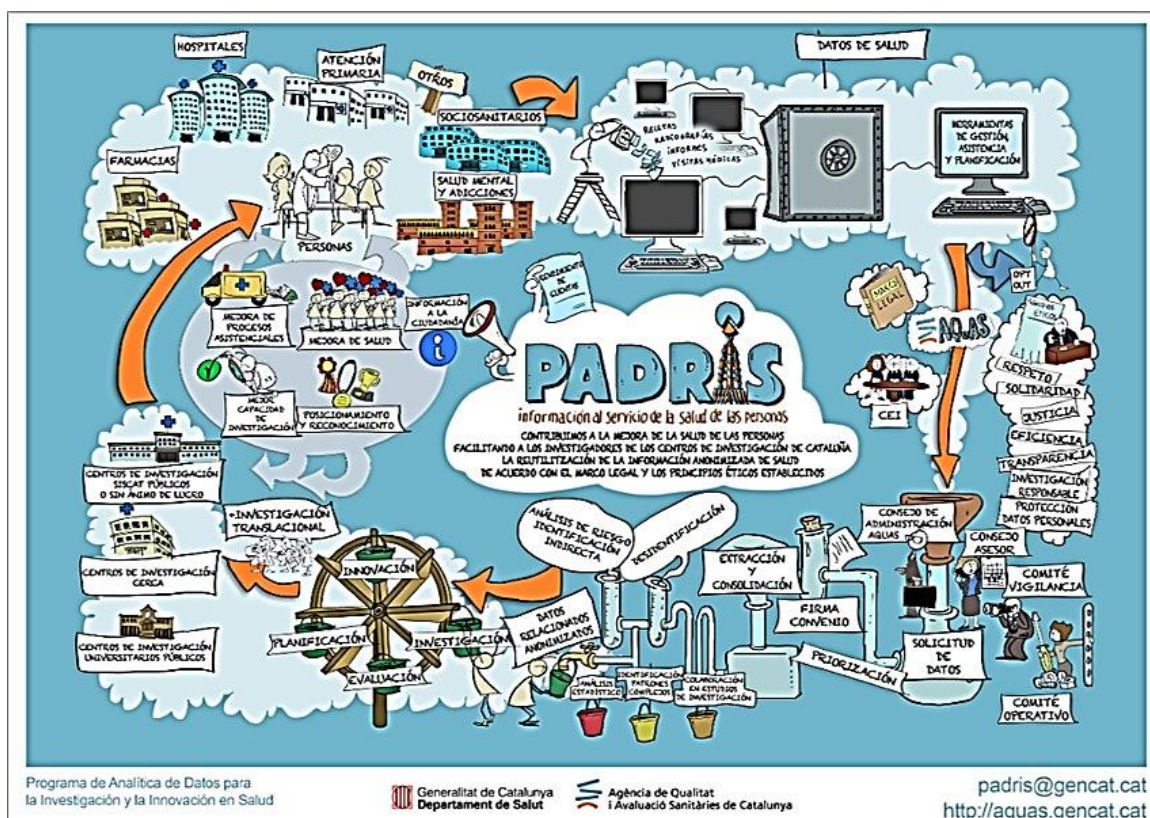


Ilustración 7. Infografía sobre el programa público PADRIS.

Fuente: Agència de Qualitat i Avaluació Sanitàries de Catalunya, 2017

Estas estas experiencias ponen de relieve lo delicado de los proyectos que impliquen el tratamiento de información sanitaria de los ciudadanos y nos plantean la necesidad no solo de cumplir con la normativa aplicable y prever las herramientas adecuadas para garantizar el respeto de los derechos de los pacientes, sino también de una correcta toma de decisiones en la gestión de los datos, de la necesidad de transparencia, información en los procesos (tanto a los profesionales, como a los pacientes) y, en definitiva, de lograr que se siga manteniendo la confianza en el sistema, confianza que es la base última de las relaciones en el ámbito sanitario.

4. PORTALES DE DATOS ABIERTOS

En años recientes, los portales de datos abiertos (PDA) o portales *open data* han sido el reflejo de una clara corriente internacional a favor de la apertura de datos en todos los niveles de las administraciones públicas. A través de los portales, las administraciones públicas ponen a disposición de los ciudadanos datos que están en su poder en formatos que facilitan la redistribución, reutilización y comercialización (Cerrillo, 2014). Dichos portales se constituyen como el instrumento fundamental que facilita tanto el aprovechamiento de datos públicos como la transparencia efectiva de la administración, considerándose como el principal producto tangible en el que se van a ver materializadas las estrategias políticas de las diferentes administraciones públicas en torno al gobierno abierto.

4.1 ¿QUÉ SON LOS PORTALES DE DATOS ABIERTOS?

Los portales de datos abiertos son plataformas digitales que sirven para almacenar, compartir, conectar y visualizar bases de datos. Son el punto de acceso a la estrategia de una organización para abrir sus datos y un punto de encuentro entre la administración, las empresas, los ciudadanos y en definitiva de todo aquel que quiera hacer uso de ellos.

Hoy en día, no existe un único modelo, pero todos los portales tienen una estructura similar con unos componentes básicos, así como una serie de componentes recomendables. Esta clasificación en obligatorios/básicos y recomendables, se ha realizado de acuerdo con los resultados de una encuesta realizada entre un colectivo de reutilizadores y responsables de portales de datos abiertos en España (FEMP, 2017):

COMPONENTES BÁSICOS
<ol style="list-style-type: none"> 1. Catálogo: es la parte central de un portal de datos abiertos y contiene un listado de todas las tablas publicadas con una descripción del contenido de la base de datos, el nombre de la agencia responsable, el número de veces que se ha visitado, la información técnica para conectarla con aplicaciones informáticas, etc. 2. Buscador simple: basado en palabras clave. 3. Filtrado: por sectores de la Notas Técnicas de Interoperabilidad (NTI), por formatos, por frecuencia de actualización, etc. 4. Condiciones de uso o términos de reutilización. 5. Medio de contacto y/o ayuda técnica. 6. Colaboración ciudadana: propuestas de nuevos conjuntos de datos para publicar. 7. Colaboración ciudadana: conjuntos de datos generados por ciudadanos/as. 8. Colaboración ciudadana: propuesta de aplicación realizada por reutilizadores/as. 9. Colaboración ciudadana: aplicaciones realizadas que han sido enviadas por reutilizadores/as. 10. Colaboración ciudadana: registro de reutilizadores (para poder recibir noticias y soporte). 11. Ejemplos de uso. 12. Servicios Interfaz de Programación de Aplicaciones (API).

Tabla 7. Componentes básicos de un portal de datos abiertos.

Fuente: FEMP (2017)

COMPONENTES RECOMENDABLES
<ol style="list-style-type: none"> 1. Servicio SPARQL¹⁹ (lenguaje estándar de consultas para la <i>web</i> semántica). 2. Herramientas de visualización.

Tabla 8. Componentes recomendables de un portal de datos abiertos.

Fuente: FEMP (2017)

¹⁹ SPARQL “es el acrónimo de Protocolo Simple y Lenguaje de Consulta de RDF. SPARQL es una tecnología que permite hacer consultas sobre información expresada en RDF, usando distintas fuentes de datos y permite obtener también los resultados en formato RDF para poder reutilizarlos en las aplicaciones” Véase <https://www.zaragoza.es/sede/portal/datos-abiertos/sparql>

El principal contenido de los portales de datos abiertos son los dataset o los conjuntos de datos, términos que hacen referencia a la categorización de los datos públicos en los catálogos de datos. Los datos en bruto se agrupan en conjuntos de datos o dataset para ser más fácilmente indexados y localizados. Según la NTI, los conjuntos de datos publicados en el catálogo deben estar documentados mediante metadatos de acuerdo a la legislación vigente, en este caso la NTI. Los campos requeridos son título, descripción detallada, la categoría a la que pertenece, el organismo responsable o publicador, y el URL de acceso. Mientras que en los campos opcionales encontraremos etiquetas, condiciones de reutilización, fecha de actualización, cobertura geográfica, etc. (NTI, 2015). En el ANEXO II se puede visualizar un ejemplo de dataset con los campos correspondientes.

Estos dataset son de temática y tipologías de formatos diversos. Con el fin de clasificar estos conjuntos de datos, la mayoría de portales tienen establecidas las categorías determinadas en la Norma Técnica de Interoperabilidad de reutilización de recursos de la información de ámbito estatal. Aunque el uso de esta taxonomía propuesta no es de obligado cumplimiento por partes de las administraciones públicas, y por ello algunos gobiernos deciden elaborar su propia descripción temática.

Dado el consenso a la hora de publicar los datos bajo las mismas categorías establecidas en la NTI, es conveniente detallar que información se están ofreciendo en cada una de ellas (PINTOS, 2012):

Ciencia y tecnología. Incluye datos sobre telecomunicación, tecnología, I+D+I, patentes, y repositorios relacionados con datos de proyectos de investigación que han sido financiados con fondos públicos (habían estado inaccesibles hasta hace unos años).

Comercio. Engloba todos aquellos repositorios que contengan información relacionada con actividades comerciales, como la información y exportación proporcionada por el ICEX, información de Cámaras de Comercio, normativa relacionada con la ordenación del comercio, registros de actividades comerciales o participantes de las mismas.

Cultura y ocio. Incluye toda la información relacionada con literatura, cine, música, danza y teatro, además de todos aquellos libros digitalizados en bibliotecas que no estén sujetos a derechos de autor.

Demografía. Información demográfica real o estadística sobre población.

Deporte. Repositorios que contengan todo tipo de información relacionada con esta temática, desde ayudas y becas a deportistas hasta listados de federaciones y asociaciones o ubicación, características e información sobre instalaciones deportivas, centros de enseñanza deportiva, etc.

Economía. Contiene almacenes de datos con información económica y financiera de todo tipo: financiación local y autonómica, estudios de evolución de precios de determinados mercados como el tecnológico, acceso a la biblioteca virtual del Ministerio de Economía y Hacienda, etc.

Educación. Contiene información tanto para estudiantes de todos los niveles y profesorado, como material que es susceptible de ser usado como material educativo.

Empleo. Se ofrece acceso a las distintas páginas de entidades públicas donde se puede consultar la información sobre las contrataciones que van a ejecutar. Además, se puede obtener información estadística sobre la población activa, datos de empleo y desempleo directamente desde el Instituto Nacional de Estadística.

Energía. Aporta información muy diversa; listados de electrodomésticos en España con su calificación de eficiencia energética, acceso a un directorio de empresas de energías renovables o acceso a un listado de los documentos aprobados por la CNE (Comisión Nacional de Energía) referente a cualquiera de las áreas posibles: gas, electricidad, energía nuclear o energía “verde”.

Hacienda. Engloba esta categoría toda la información relacionada con la Agencia Tributaria, desde impuestos y recaudación hasta la consulta sobre la doctrina tributaria, consulta de bienes enajenados por la Agencia Tributaria incluyendo las fechas de las próximas subastas, etc.

Industria. Almacenes de datos relacionados con la actividad empresarial e industrial, tanto a nivel de pymes como grandes empresas. Buscador de prospecciones

tecnológicas industriales cubriendo sectores tan diversos como la automoción, la electrónica, el sector del metal, etc.

Legislación y justicia. Se tiene acceso a todo tipo de jurisprudencia histórica y actualizada de todos los niveles: sentencias del Tribunal Constitucional, Audiencia Nacional, Tribunal Superior de Justicia, leyes, decretos, etc.; bases de datos de legislación extranjera, propuestas legislativas, publicaciones del Ministerio de Justicia, etc.

Medioambiente. Ofrece información sobre las diversas confederaciones hidrográficas, umbrales de información climatológica, boletines hidrológicos o bases de datos de ayuda a los sectores agroalimentarios, pesquero y medioambiental; información sobre el Instituto Español de Oceanografía, etc.

Medio rural y pesca. Incluye entre otros temas, almacenes relacionados con sanidad vegetal y animal, información sobre producciones agrícolas y ganaderas; información sobre regulación de mercados como la información sobre los precios percibidos, pagados, salarios agrarios; información sobre etiquetado de producción, denominaciones de origen e indicaciones geográficas protegidas, etc.

Salud. Aquí encontramos repositorios con información diversa. Un repositorio básico es el banco de datos del Ministerio de Sanidad, Servicios Sociales e Igualdad con información sobre el barómetro sanitario, datos de la Encuesta Nacional de Salud, diversas estadísticas y el catálogo de centros de atención primaria del Servicio Nacional de Salud, así como el catálogo de hospitales. También encontramos información sobre salud ambiental, salud mental, enfermedades raras.

Sector público. Es la categoría donde más información diversa podemos encontrar: organigramas de organismos, presupuestos, contrataciones y licitaciones de éstos. Del mismo modo se encuentra información sobre boletines oficiales y normativa relacionada con el ejercicio de la función pública. También hay información interesante como la recopilación de la presencia en las redes sociales de la Administración General del Estado.

Seguridad. Es la categoría más limitada. Aquí tendremos acceso a la localización de comisarías de policía, la búsqueda de plazas en las fuerzas armadas, buscador de cartas

náuticas del Instituto Hidrográfico de la Marina o almacenes relacionados con las instituciones penitenciarias.

Sociedad y bienestar. Recoge todos los repositorios relacionados con aquellos segmentos de la sociedad relacionados con la cultura del bienestar: mayores, niños, sectores desfavorecidos, discapacitados, etc.

Transporte. Información sobre el control del tráfico en distintos medios, estado de las carreteras y vías circulatorias, y siniestralidad. No solo aplicable al segmento del automóvil sino de la aviación civil, transporte de mercancías, y marina mercante. Localizaciones ITV o Jefaturas Provinciales de Tráfico, consulta sobre el saldo de los puntos del carné de conducir, etc.

Turismo. Ofrece información referente a la ocupación hotelera a través de buscadores de alojamientos turísticos. También sobre oficinas exteriores de información turística, comportamientos y hábitos turísticos, etc.

Urbanismo e infraestructuras. Aquí tenemos acceso a los programas como Urbanismo en Red de la entidad pública Red.es, ofical virtual del catastro o Instituto Geográfico Nacional. Información cartográfica, mapas georreferenciados, información catastral, etc.

Vivienda. Refiere a todos los almacenes de datos relacionados con el sector de la construcción y el sector inmobiliario.

SECTORES
Ciencia y tecnología Incluye: innovación, investigación, I + D + I, telecomunicaciones, internet, y sociedad de la información
Comercio Incluye: consumo
Cultura y ocio Incluye: tiempo libre
Demografía Incluye: inmigración y emigración, familia, mujeres, infancia, personas mayores, y padrón
Deporte Incluye: instalaciones deportivas, federaciones, y competiciones
Economía Incluye: deuda, moneda, banca, y finanzas
Educación Incluye: escuelas, educación, y actividades de formación
Empleo Incluye: formación
Energía Incluye: energías renovables
Hacienda Incluye: impuestos
Industria Incluye: minería
Legislación y justicia Incluye: registros
Medio ambiente Incluye: meteorología, geografía, y conservación de la fauna y flora
Medio rural Incluye: agricultura, ganadería, pesca, y silvicultura
Salud Incluye: sanidad
Sector público Incluye: presupuestos, organigrama institucional, legislación interna, y función pública
Seguridad Incluye: protección civil
Sociedad y bienestar Incluye: participación ciudadana, marginación, envejecimiento activo, dependencia, jubilación, seguros y pensiones, prestaciones, y subvenciones
Transporte Incluye: transporte, comunicaciones, y tráfico
Turismo Incluye: alojamiento, hostelería, y gastronomía
Urbanismo e infraestructuras Incluye: construcción, infraestructuras, equipamientos públicos, y saneamiento público
Vivienda Incluye: mercado inmobiliario y vivienda

Tabla 9. Cuadro resumen: categorías determinadas en la NTI

Fuente: NTI (2013).

4.2 PORTAL DE DATOS ABIERTOS DE LA UNIÓN EUROPEA

Los movimientos alrededor de los datos abiertos en el sector público, que se producen desde el año 2009 en Europa, han fomentado la creación de múltiples portales en el territorio de la Unión Europea. El portal de datos abiertos de la Unión Europea²⁰ constituye un punto de acceso único para una gran diversidad de datos que han sido producidos por instituciones y otros organismos de la Unión Europea (Arquero y Marco, 2015). Es considerado por ello un importante activo de información para el conjunto de estados miembros de la Unión Europea y, por consiguiente, un importante vehículo de propagación de la información que la Unión Europea genera y pone a disposición pública de la sociedad en general y de los investigadores. Con este recurso de acceso fácil y gratuito a los datos, la Comisión Europea pretende fomentar el uso público de los datos y su reutilización en beneficio de nuevos proyectos innovadores y transfronterizos.

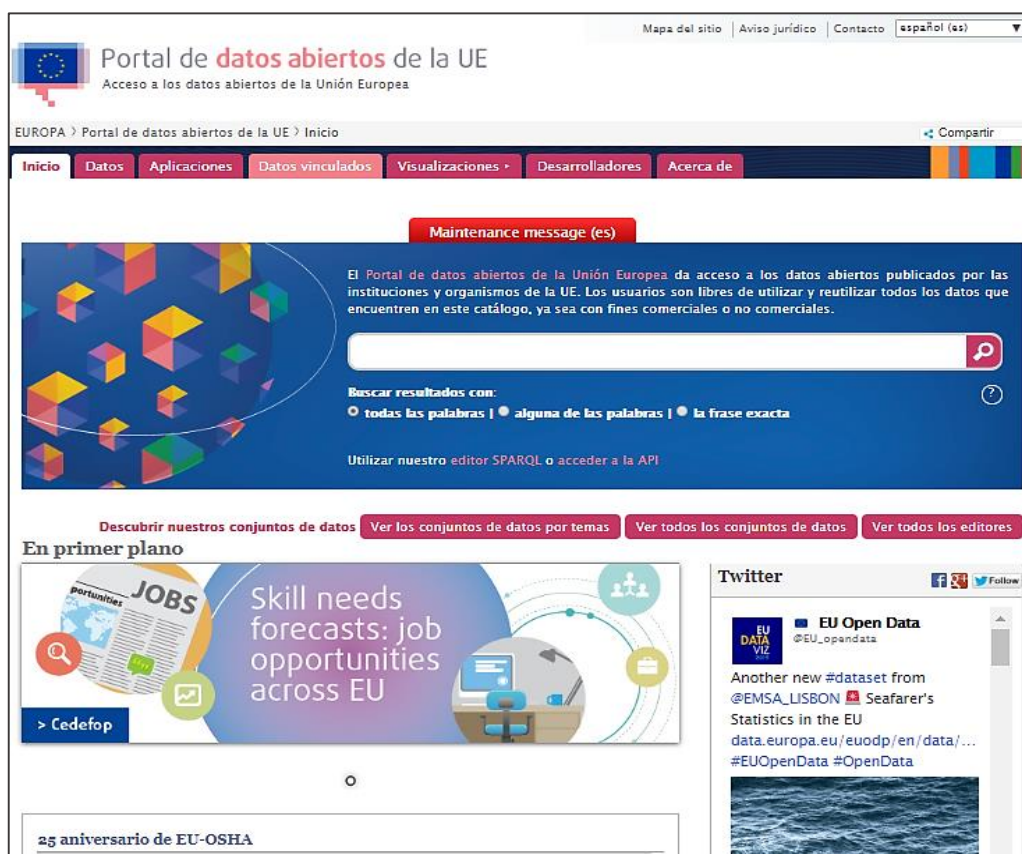


Ilustración 8. Captura de la página de inicio del portal de datos abiertos de la Unión Europea

²⁰ Puedes acceder a él a través de <https://data.europa.eu/euodp/es/home>

El portal de datos abiertos de la Unión Europea (Open Data Europe Portal en inglés) se enlaza con la Directiva 2013/37/UE, del Parlamento Europeo y del Consejo, relativa a la reutilización de la información en el sector público y se relaciona con la Agenda Digital Europea, iniciativa emprendida por la Comisión Europea (2010) que, promueve la comunicación de la Comisión Europea sobre datos abiertos como motor para la innovación, el crecimiento y la gobernanza transparente.

Los datos que se facilitan a través del portal pueden utilizarse y reutilizarse con fines comerciales o no comerciales, con lo cual se busca impulsar su potencial económico. Su reutilización, de forma gratuita, conlleva únicamente la citación de la fuente. Toda la información que se ofrece está sujeta a una cláusula de exención de responsabilidad, a un aviso de derechos de autor y a las normas aplicables de protección de datos personales (Arquero y Marco, 2015).

El portal, disponible en 24 idiomas, constituye también una fuente de importante valor para la comparación de datos entre los diversos organismos productores y fomenta el establecimiento de relaciones entre éstos y la comunidad de usuarios. Además, como herramienta, dispone de un área exclusiva para proveedores de datos. En cuanto a los datos, el portal presenta un catálogo de metadatos que facilita su acceso. Los metadatos son vinculados y normalizados basándose en estándares que facilitan su recuperación y reutilización posterior. Estos conjuntos de datos se encuentran disponibles en formatos que permiten su interpretación humana, y son también legibles por un ordenador para facilitar su inmediata reutilización. El total de recursos o conjuntos de datos disponibles es de 13.944²¹ (Arquero y Marco, 2015). De los cuales 2.654 pertenecen a la temática de Salud.

Cabe recordar que en Europa también existe el European Data Portal (EDP)²², el cual no debe ser confundido con el Open Data Europe Portal (ODP). La diferencia entre ambos reside en que el ODP es el portal de datos abiertos de la Unión Europea que contiene conjuntos de datos recopilados y publicados por las instituciones europeas mientras que

²¹ Datos oficiales ofrecidos por el Portal de datos abiertos de la UE.

²² <https://www.europeandataportal.eu/>

el EDP es un portal europeo que recoge metadatos de portales del sector público en toda Europa. Por lo tanto, el EDP se centra en los datos puestos a disposición por los países europeos y además también cosecha metadatos de ODP. La Comisión Europea está explorando actualmente cómo acercar esos dos portales.

4.3 PORTAL DE DATOS ABIERTOS DE ESPAÑA

Como señala García García (2010), la mera existencia de normas no es suficiente para desarrollar un adecuado gobierno digital, por lo que el Estado español ha puesto en marcha varias iniciativas para acompañar las recientes regulaciones legales. Una de esas iniciativas propulsoras, es el Proyecto Aporta, impulsor de la RISP cuyo cometido es el de apoyar la Ley 37/2007 sobre Reutilización de la información del sector público.

4.3.1 PROYECTO APORTA

Aporta, nació en 2009 con tres objetivos básicos: fomentar una cultura favorable a la apertura de datos públicos, facilitar que las AAPP acometan dicha apertura e impulsar el mercado de la RISP (Curto, 2017). Y es promovido por el Ministerio de Economía y Empresa, el Ministerio de Política Territorial y Función Pública y la Entidad Pública Empresarial Red.es.

El Proyecto Aporta es consciente de la importancia y el valor de la información de la información pública, por lo que va a realizar actividades de concienciación de su valor, así como otras que pretenden estimular la reutilización de dicha información. Para ello se crea el portal [www.aporta.es](http://datos.gob.es) (desde julio de 2012, <http://datos.gob.es>) y se realizan acciones específicas como talleres, seminarios y encuentros, elaboración de materiales formativos, etc. (Curto, 2017).



Ilustración 9. Captura de pantalla de la página de inicio del portal de datos abiertos de España.

Fuente: datos.gob.es

La Iniciativa Aporta y datos.gob.es están dirigidos a todos los agentes del ecosistema de los datos:

- Usuarios, ciudadanos o profesionales que demandan datos y/o quieren conocer novedades, aplicaciones o servicios relacionados con ellos.
- Organismos públicos que proporcionan y utilizan datos públicos, y que quieren estar al tanto de novedades del sector.
- Reutilizadores e infomediarios que precisan fuentes de datos para crear productos y servicios que quieren dar a conocer.

Esta iniciativa se desarrolla en torno siete líneas de actuación²³:



Ilustración 10. Líneas de actuación del Proyecto aporta.

Fuente: datos.gob.es

²³ Para profundizar en las líneas de actuación del Proyecto Aporta consulte <https://datos.gob.es/es>

4.3.2 CATÁLOGO DE INFORMACIÓN PÚBLICA DEL SECTOR PÚBLICO

El portal organiza y gestiona el Catálogo de Información Pública del Sector Público. El objetivo del catálogo es facilitar a los ciudadanos y empresas la identificación y acceso de los conjuntos de datos disponibles en el portal datos.gob.es para favorecer, de esta manera, la reutilización de dicha información.

Actualmente el catálogo cuenta con un total de 24.753 conjuntos de datos distribuidos entre las categorías descritas en la NTI y detalladas anteriormente:

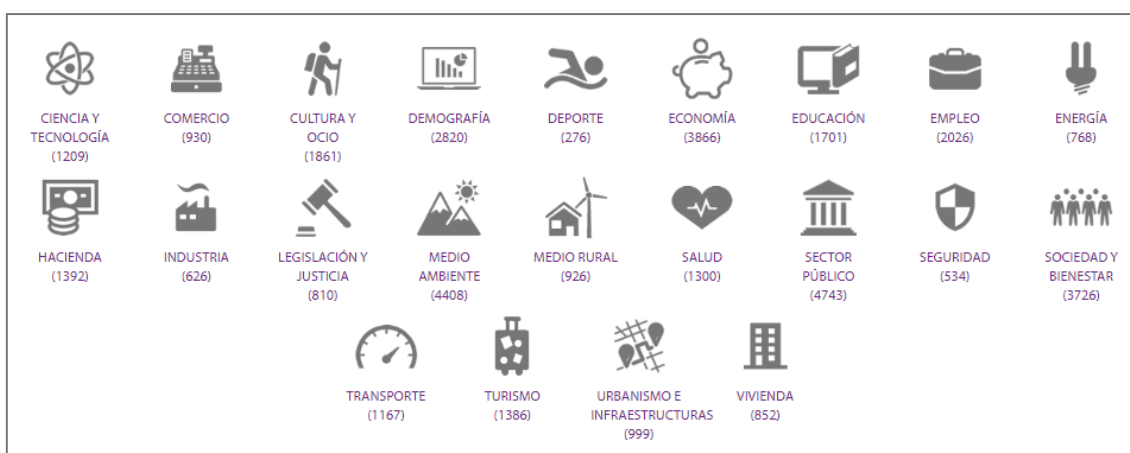


Ilustración 11. Captura de pantalla de las temáticas establecidas en el portal datos.gob.es

Fuente: datos.gob.es

Entre la diversa información que el catálogo ofrece podemos encontrar y filtrar por:

- Total, de conjuntos de datos por formatos
- Organismos publicadores más el total de conjuntos publicados por el mismo.
- Total, de datos publicados según el nivel de la Administración
- Frecuencia de actualización de los dataset
- Etiquetas más utilizadas

Además, cuenta con un buscador para búsquedas personalizadas y permite ordenar los resultados en función de nuestras necesidades.

El Catálogo del portal datos.gob.es fue galardonado por la plataforma LAPSI en enero de 2012 como el portal más amigable de información pública de la Unión Europea.

4.3.3 MAPA DE INICIATIVAS DATOS.GOB.ES

Aunque son muchos apartados de los que dispone el portal, merece una mención especial el Mapa de Iniciativas que ponen a disposición de los usuarios. Esta herramienta permite visualizar de forma más sencilla cual es el impacto que está teniendo la apertura de datos a nivel estatal, autonómico, local y también de Universidad, al dibujar un mapa con las iniciativas de apertura de datos en España, tal y como se muestra en la siguiente ilustración:

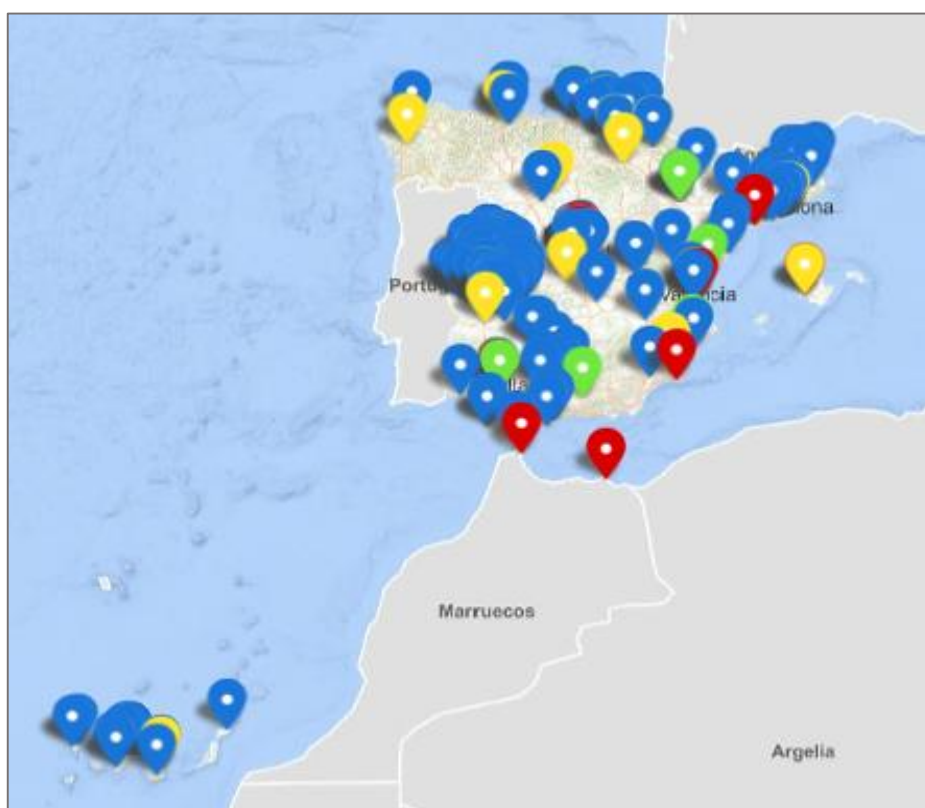


Ilustración 12. Mapa de iniciativas en España

Fuente: datos.gob.es

Los puntos de localización rojos, hacen referencia a las iniciativas estatales; los puntos amarillos, a las iniciativas autonómicas; los puntos azules, a las iniciativas locales, y los puntos verdes, a las iniciativas de las Universidades.

Además, este Mapa de Iniciativas no solo permite hacer un filtrado por iniciativas de apertura de datos por nivel de administración, sino también por temática y grado de

reutilización. Y ha sido la herramienta base para el acceso a los portales web de datos abiertos de las comunidades autónomas que han sido objeto de estudio.

4.4 PORTALES DE DATOS ABIERTOS DE LAS COMUNIDADES AUTÓNOMAS ESPAÑOLAS

Con la aparición de leyes anteriormente mencionadas, sobre todo con la LRISP, los gobiernos de las comunidades autónomas españolas empezaron a comprender la necesidad de abrir sus datos y por tanto la necesidad de crear portales de datos abiertos para recogerlos. Si bien es cierto que, desde la aprobación de estas leyes, se han visto incrementados los proyectos de datos abiertos en España, a día de hoy, ciudades como Ceuta y Melilla no disponen de portal de datos abiertos.

A continuación, se detallará sin entrar en profundidad la información que ofrece datos.gob.es de cada portal de datos abiertos por CCAA:

ANDALUCÍA	
Nombre	Portal Datos Abiertos Junta de Andalucía (Junta de Andalucía)
Enlace Web	https://www.juntadeandalucia.es/datosabiertos/portal.html
Tipo de buscador	Sencillo
Licencia	Creative Commons Reconocimiento (CC BY)
Otros enlaces	Canal de solicitud Recursos formativos

Tabla 10. Información básica aportada por datos.gob.es para el portal de Andalucía

ARAGÓN	
Nombre	Aragón Open Data
Enlace Web	https://opendata.aragon.es/
Tipo de buscador	Avanzado
Licencia	Creative Commons Reconocimiento (CC BY)
Otros enlaces	Canal de solicitud Catálogo de aplicaciones Punto SPARQL: Punto SPARQL API Ckan Recursos formativos Colabora con datos.gob.es

Tabla 11. Información básica aportada por datos.gob.es para el portal de Aragón

ASTURIAS	
Nombre	Open data del Principado de Asturias
Enlace Web	(Mirar bibliografía)
Tipo de buscador	Avanzado
Licencia	Creative Commons Zero (CC0)

Tabla 12. Información básica aportada por datos.gob.es para el portal de Asturias

ISLAS BALEARES	
Nombre	Open Data GOIB
Enlace Web	http://www.caib.es/sites/opendatacaib/es/inicio-27248/?campa=yes
Tipo de buscador	Sencillo
Licencia	http://www.caib.es/sites/opendatacaib/es/politica i termes dus/
Otros enlaces	Restful API Colabora con datos.gob.es

Tabla 13. Información básica aportada por datos.gob.es para el portal de las Islas Baleares

ISLAS CANARIAS	
Nombre	Open Data del Gobierno de Canarias
Enlace Web	https://opendata.gobiernodecanarias.org/opendata/
Tipo de buscador	Avanzado
Licencia	Creative Commons Zero (CC0)
Otros enlaces	Canal de solicitud Catálogo de aplicaciones Recursos formativos Colabora con datos.gob.es

Tabla 14. Información básica aportada por datos.gob.es para el portal de las Islas Canarias

CANTABRIA	
Nombre	Datos Abiertos del ICANE
Enlace Web	https://datos.icane.es/
Tipo de buscador	-
Licencia	http://www.icane.es/footer/legal-advisory
Otros enlaces	Punto SPARQL: Punto SPARQL Web Service

Tabla 15. Información básica aportada por datos.gob.es para el portal de Cantabria

CASTILLA LA MANCHA	
Nombre	Datos Abiertos CLM (Junta de Castilla La Mancha)
Enlace Web	https://datosabiertos.castillalamancha.es/
Tipo de buscador	Avanzado
Licencia	Creative Commons - Reconocimiento (CC-by 4.0)
Otros enlaces	Canal de solicitud Colabora con datos.gob.es

Tabla 16. Información básica aportada por datos.gob.es para el portal de Castilla La Mancha

CASTILLA Y LEÓN	
Nombre	Portal de Datos Abiertos Junta de Castilla y León
Enlace Web	https://datosabiertos.castillalamancha.es/
Tipo de buscador	Avanzado
Licencia	Creative Commons Reconocimiento (CC BY)
Otros enlaces	Estrategia apertura de datos Canal de solicitud Catálogo de aplicaciones Recursos formativos Colabora con datos.gob.es

Tabla 17. Información básica aportada por datos.gob.es para el portal de Castilla y León

CATALUÑA	
Nombre	Portal de Datos Abiertos Generalitat de Catalunya
Enlace Web	http://governobert.gencat.cat/ca/dades_obertes/
Tipo de buscador	Avanzado
Licencia	Creative Commons Reconocimiento-SinObraDerivada (CC BY-ND), Creative Commons Reconocimiento (CC BY), Licencia Abierta
Otros enlaces	Canal de solicitud Recursos formativos

Tabla 18. Información básica aportada por datos.gob.es para el portal de Cataluña

COMUNIDAD VALENCIANA	
Nombre	Dades Obertes de la Generalitat Valenciana
Enlace Web	http://www.dadesobertes.gva.es/va/dataset?q=&sort=metadata_modified+desc

Tabla 19. Información básica aportada por datos.gob.es para el portal de la Comunidad Valenciana

EXTREMADURA	
Nombre	Portal de Transparencia y Participación ciudadana de Extremadura
Enlace Web	http://gobiernoabierto.juntaex.es/
Tipo de buscador	Avanzado
Licencia	-
Otros enlaces	Descarga del catálogo Canal de solicitud API Ckan

Tabla 20. Información básica aportada por datos.gob.es para el portal de Extremadura

GALICIA	
Nombre	Portal Open Data de la Xunta de Galicia
Enlace Web	http://abertos.xunta.es/portada
Tipo de buscador	Avanzado
Licencia	Creative Commons Zero (CC0)
Otros enlaces	Descarga del catalogo Canal de solicitud Catálogo de aplicaciones Recursos formativos Colabora con datos.gob.es

Tabla 21. Información básica aportada por datos.gob.es para el portal de la Xunta de Galicia

LA RIOJA	
Nombre	Dato Abierto La Rioja
Enlace Web	https://web.larioja.org/dato-abierto
Tipo de buscador	Avanzado
Licencia	Creative Commons Reconocimiento (CC BY), Creative Commons Reconocimiento-NoComercial (CC BY-NC)
Otros enlaces	Canal de solicitud Colabora con datos.gob.es

Tabla 22. Información básica aportada por datos.gob.es para el portal de La Rioja

COMUNIDAD DE MADRID	
Nombre	Portal de Datos Abiertos de la Comunidad de Madrid
Enlace Web	http://www.comunidad.madrid/gobierno/datos-abiertos
Tipo de buscador	Avanzado
Licencia	http://www.comunidad.madrid/gobierno/datos-abiertos/reutiliza#condiciones-uso
Otros enlaces	Canal de solicitud Restful API

Tabla 23. Información básica aportada por datos.gob.es para el portal de la Comunidad de Madrid

REGIÓN DE MURCIA	
Nombre	Datos Abiertos en la Región de Murcia
Enlace Web	http://datosabiertos.regiondemurcia.es/
Tipo de buscador	Sencillo
Licencia	http://datosabiertos.regiondemurcia.es/avisolegal
Otros enlaces	Canal de solicitud Catálogo de aplicaciones API Ckan

Tabla 24. Información básica aportada por datos.gob.es para el portal de la Región de Murcia

COMUNIDAD FORAL DE NAVARRA	
Nombre	Portal de Gobierno abierto de la Comunidad Foral de Navarra
Enlace Web	http://www.gobiernoabierto.navarra.es/es/open-data
Tipo de buscador	Avanzado
Licencia	Creative Commons Zero (CC0)
Otros enlaces	Canal de solicitud Catálogo de aplicaciones Recursos formativos

Tabla 25. Información básica aportada por datos.gob.es para el portal de Navarra

PAIS VASCO	
Nombre	Open data Euskadi
Enlace Web	http://opendata.euskadi.eus/inicio/
Tipo de buscador	Avanzado
Licencia	Creative Commons Reconocimiento (CC BY)
Otros enlaces	Canal de solicitud Catálogo de aplicaciones Punto SPARQL: Punto SPARQL Recursos formativos

Tabla 26. Información básica aportada por datos.gob.es para el portal del País Vasco

5. RESULTADOS

Para una comprensión más sencilla y eficaz de los resultados, éstos se presentarán por tipología de indicadores. Todas las ilustraciones mostradas son de elaboración propia.

5.1 INDICADORES SOBRE LOS PORTALES

En este apartado encontraremos tres indicadores, dos mas generales y uno mas específico. Los tres ayudarán a la comprensión de este estudio. El primero de todos hace referencia al *total de CCAA que disponen de portales de datos abiertos*.

Tal y como muestra la ilustración 13, un 89% de CCAA disponen de portal de datos abiertos mientras que tan solo un 11% no disponen aún de ello. Este bajo porcentaje

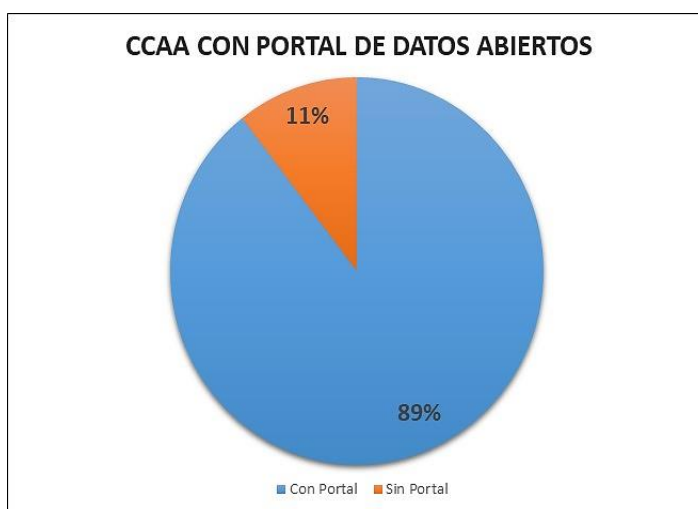


Ilustración 13. CCAA con Portal de Datos Abiertos

corresponde a las Comunidades Autónomas de Ceuta y Melilla. Se ha comprobado que Melilla, aunque cuenta con un enlace web en el datos.gob.es, éste da error constantemente por lo que ha sido descartado como válido para este estudio.

El segundo indicador corresponde a *cuantos portales de datos abiertos publican sus conjuntos de datos bajo la temática propuesta por la NTI*.

Como puede observarse en la ilustración 13, un 80% de las CCAA siguen la propuesta de la NTI, lo que supone más de la tercera parte del total. Después hay un 13% de CCAA que organizan sus datos siguiendo unos criterios propios. Este porcentaje corresponde a las CCAA de Galicia y Cantabria, que tienen su propia taxonomía.

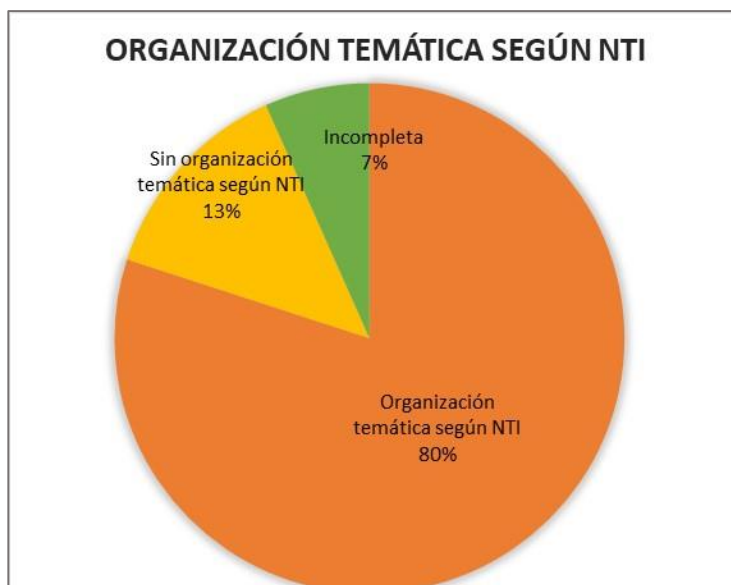


Ilustración 14. PDA organizados según la temática de NTI

Y por último tenemos un 7% que, aunque hacen un intento por seguir la temática NTI, ésta no es del todo completa y algunos conjuntos están bajo otras etiquetas diferentes. Como, por ejemplo, es el caso de Asturias.

Y el tercer indicador de este apartado es *el número total de portales consultados frente al número total de portales con conjuntos de datos en Salud*.

En este caso, todos los PDA de las CCAA cuentan con conjuntos de datos en salud, por lo que el resultado obtenido en este indicador es de un 100%. Este dato es significativo si lo vemos desde el punto de vista de la importancia de publicar datos en salud, ya que es un tema social de preocupación.

5.2 INDICADORES DE CONJUNTO DE DATOS O DATASET

En este apartado encontraremos toda la información relacionada con el análisis de los conjuntos de datos. Se pretende averiguar cuanto hay publicado actualmente en los PDA, y cuanto de este total corresponde a datos sanitarios.

El primer indicador es *total de conjuntos de datos abiertos por CCAA*. Nos permitirá conocer que CCAA tienen más dataset publicados en sus portales.

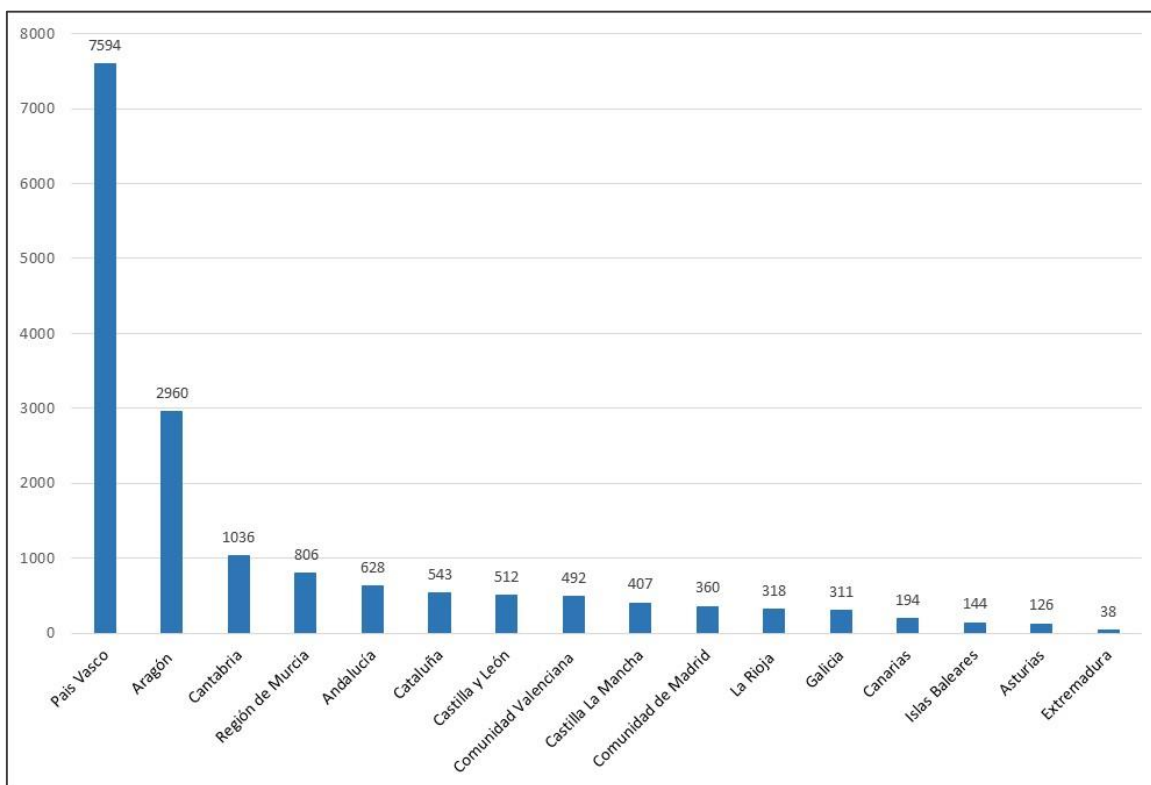


Ilustración 15. Dataset publicados por CCAA

Se puede observar una diferencia abismal entre la CCAA que más publica frente a la que menos publica. El País Vasco cuenta con un total de 7594 conjuntos de datos publicados, mientras que Extremadura tan solo cuenta con 38. Tan solo tres CCAA superan los 1000 dataset publicados: País Vasco que despunta con diferencia, seguida por Aragón con 2960 conjuntos de datos y en tercer lugar Cantabria con 1036 conjuntos de datos. El resto de CCAA tienen niveles de publicación inferiores.

El segundo y el tercer indicador están relacionados. El segundo viene determinado por la *cantidad de dataset publicados en salud por CCAA* e indicará las CCAA que más esfuerzos hacen en publicación de datos sanitarios.

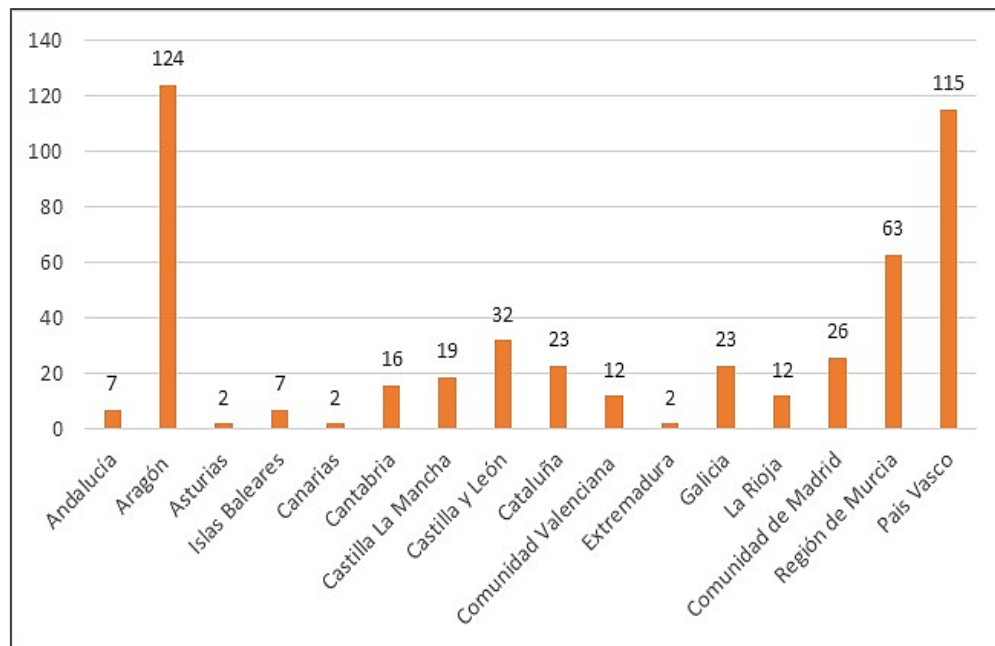


Ilustración 16. Conjuntos de datos publicados en Salud por CCAA

En este caso, País Vasco sigue despuntando con 115 conjuntos, pero ha sido superado por Aragón que cuenta con 124 dataset. El gráfico muestra claramente cierta heterogeneidad a la hora de publicar datos en Salud, pero la tendencia es clara: se publica poco en general. Tres CCAA tan solo tienen dos conjuntos de datos en salud publicados: Asturias, Canarias y Extremadura.

Tercer indicador. Si comparamos *el total de datos publicados frente al total de datos publicados en salud* obtendremos el porcentaje que representa este último en el porcentaje total de datos. Una cifra a priori, muy insignificante respecto al total. Tan solo 2,9% (485dataset) corresponde a dicha temática. El 97,1% restante es formado por el resto de categorías, equivalente a 16.087 dataset.

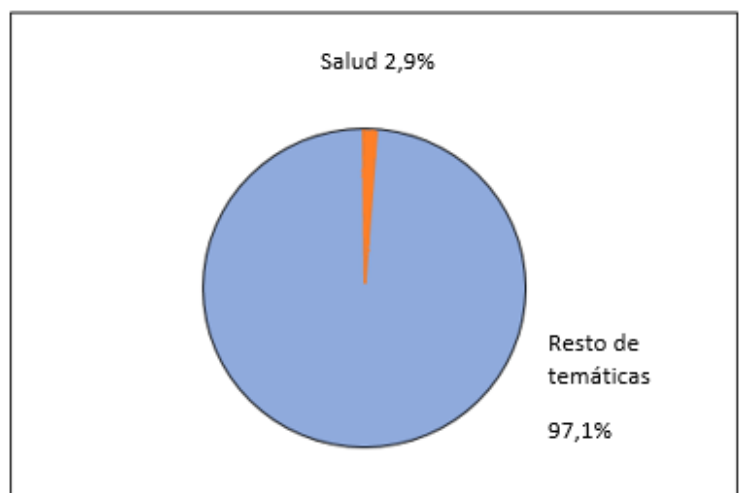


Ilustración 17. Porcentaje representativo de datos en salud en el total de datos publicados

A continuación, en la ilustración 18 podemos ver cómo se distribuyen los porcentajes por temáticas. Hay dos categorías que despuntan con respecto al resto: sector público con un 14% (2308 dataset) y economía con un 12% (1937 dataset) respectivamente. Le siguen de cerca medio ambiente y demografía con un 9% (1539

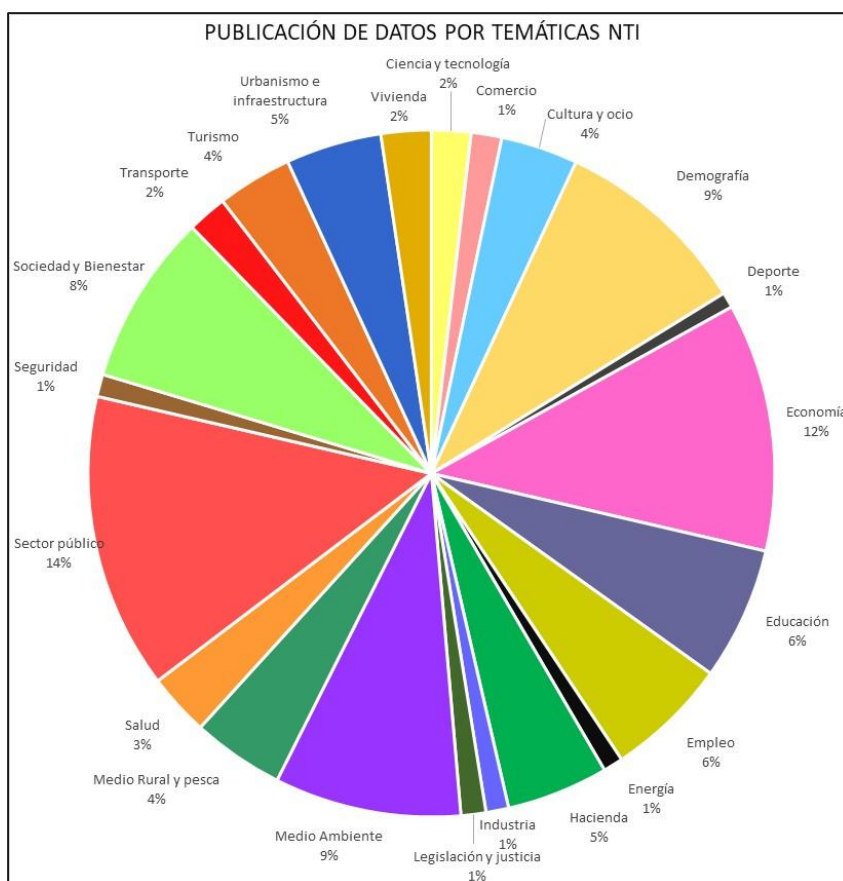


Ilustración 18. Publicación de datos por temáticas NTI

dataset) y sociedad y bienestar con un 8% (1336 dataset). Es significativo resaltar que hasta cinco temáticas cuentan con tan solo un 1% de representación (entre 100 y 250 dataset). El resto oscilan entre el 2- 6%, es decir, entre 200 y 600 dataset publicados.

En un análisis más detallado por saber que se está publicado en Salud, se ha procedido a generar un *indicador por materias especializadas dentro de la categoría de salud*. La clasificación ha sido generada y normalizada por la persona que confecciona esta investigación bajo criterios propios y por experiencia laboral en el sector sanitario. El resultado ha sido el siguiente:

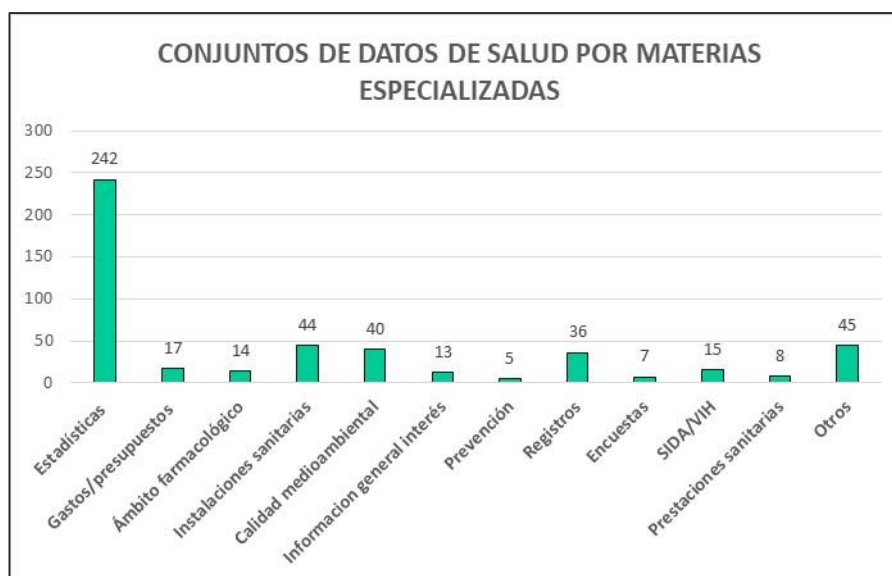


Ilustración 19. Conjuntos de datos de salud por temas especializados

La ilustración 18 muestra claramente dos hechos: el primero, la clara tendencia por parte de las AAPP a publicar datos en forma de estadísticas sobre diversos temas, y la segunda, la tendencia a publicar sobre instalaciones sanitarias o infraestructuras. Dos materias cuyos datos son de más fácil acceso que el resto por su baja privacidad. Hay que resaltar en este caso que materias como *prevención*, *encuestas* o *prestaciones* sanitarias apenas cuentan con 5, 7 y 8 conjuntos de datos respectivamente. Dado que el análisis es sobre 16 CCAA, es un dato bastante significativo.

Como se ha mencionado anteriormente, existe una gran diferencia entre la cantidad de dataset cuya temática son estadísticas con el resto de temáticas específicas que se ha creído conveniente proceder a detallar dicha temática en submaterias. El siguiente gráfico muestra cual ha sido el resultado:

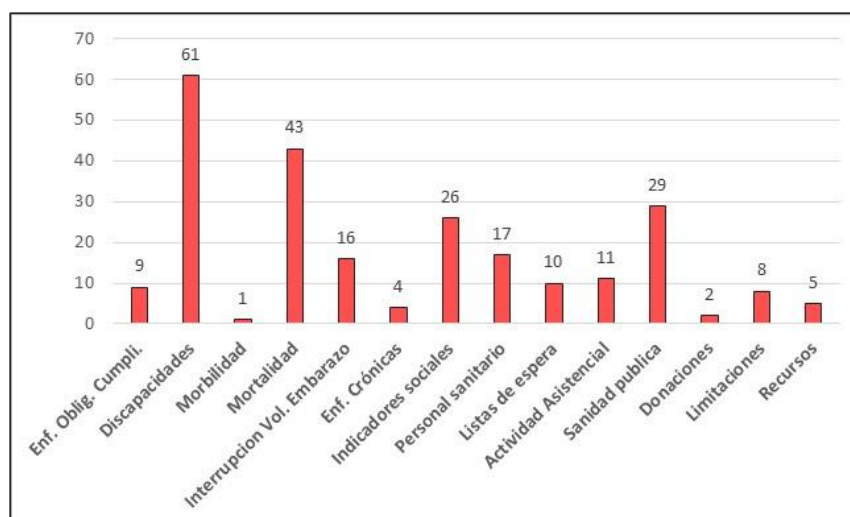


Ilustración 20. Desglose de la materia específica Estadísticas en submaterias.

La submateria sobre la que más se publican datos sanitarios es discapacidades, con un total de 61 conjuntos de datos al respecto. Le sigue Mortalidad con 43 conjuntos de datos y sanidad pública con 29. Morbilidad y Donaciones aportan un porcentaje poco significativo. Se puede extraer de este gráfico que nueve de las catorce submaterias hacen referencia a estadísticas demográficas sanitarias.

5.3 INDICADORES DE CALIDAD DE LOS CONJUNTOS DE DATOS EN SALUD

Para comprobar la calidad de los conjuntos de datos publicados en salud, nos basaremos en los formatos utilizados para la publicación de éstos. Como hemos indicado anteriormente, no todos los formatos tienen el mismo nivel de reutilización. Los indicadores que a continuación se exponen pretenden clarificar cual es la calidad de los datos sanitarios publicados actualmente en los PDA de las CCAA españolas, y algunos de ellos los analizaremos de acuerdo con el esquema de cinco estrellas de Berners-Lee.

Hay que puntualizar, que un mismo conjunto de datos ha podido ser publicado en varios formatos diferentes, por lo que el total de conjuntos de datos no va a coincidir con los totales de conjuntos de datos de los formatos.

TOTAL DATOS PUBLICADOS EN SALUD	TOTAL DATOS PUBLICADOS EN SALUD SEGÚN LOS FORMATOS UTILIZADOS.
485	833

Tabla 27. Comparativa de datos publicados en salud

Dicho esto, el primero indicador que analizamos en este apartado muestra el *total de conjuntos de datos publicados en salud por formatos*. Será interesante conocer cuáles son los formatos más utilizados a la hora de publicar en Salud.



Ilustración 21. *Formatos utilizados en la publicación de datos de salud*

En este caso, el formato CSV es con diferencia el formato más utilizado para publicar datos en salud, con un total de 306 dataset. Le siguen los formatos XLS o XLSX, con un total de 195 dataset. Y muy por debajo los formatos JSON/geoJSON con 82 conjuntos y XML con 66 respectivamente. Es impactante comprobar las diferencias que existen entre la utilización de unos formatos u otros.

El segundo indicador, que está relacionado con el primero, hace referencia a *qué comunidades autónomas están publicando en los formatos más reutilizables*.

La siguiente tabla muestra los formatos según el modelo de Berners-Lee por CCAA. Aquí podremos ver como todas las comunidades autónomas tienen datos publicados en nivel 3 estrellas, es decir, los datos que publican en sus webs tienen formato estructurado y, además, son abiertos y no propietarios. Cualquier persona puede acceder y utilizar los datos estos datos de forma sencilla. También es frecuente utilizar formatos de 2 estrellas, es decir, formatos estructurados (legibles por máquinas) que facilitan su tratamiento pero que están en formato propietario, lo que puede exigir la compra de alguna licencia para poder trabajar con la información. La práctica más habitual entre CCAA empieza a ser la publicación de conjuntos de datos en estos dos formatos. Aunque quedan CCAA como Andalucía, Castilla La Mancha, Valencia o Galicia, que siguen usando

formatos de nivel 1. Formatos no estructurados que requieren de la aplicación de técnicas y esfuerzo extra, y que por lo tanto no son recomendables utilizar.

Cabe señalar destacar de esta tabla a la Comunidad autónoma de Madrid que, a pesar de utilizar solo dos formatos distintos para la publicación de sus datos, estos son de nivel 3 estrellas. Es decir, no utiliza demasiados formatos, pero hace un esfuerzo considerable en publicar los conjuntos de datos en formatos de mas fácil acceso y reutilización.

	FORMATOS 1 ESTRELLA	FORMATOS 2 ESTRELLAS	FORMATOS 3 ESTRELLAS	FORMATOS 4 ESTRELLAS
ANDALUCÍA	PDF	SHP; XLS	CSV; XML; ODS; JSON; HTML	RDF
ARAGÓN		XLS	CSV	
ASTURIAS			XML	
ISLAS BALEARES		XLS; SHP	CSV; KML; JSON/GEOJSON	
CANARIAS		SHP	CSV; GEOJSON	
CANTABRIA		XLS	HTML; JSON	RDF
CASTILLA LA MANCHA	PDF	XLS/XLSX	CSV; HTML;	
CASTILLA Y LEÓN			CSV; XML	
CATALUÑA		SHP	CSV; XML; RSS; KML; HTML; JSON	RDF
COMUNIDAD VALENCIANA	PDF		CSV; JSON; XML	
EXTREMADURA		XLS	CSV	
GALICIA	ZIP	XLS	CSV; KML; HTML; WMS; WFS; RSS	
LA RIOJA		XLS	CSV; XML; JSON	
COMUNIDAD DE MADRID			CSV; JSON	
REGIÓN DE MURCIA		XLS	CSV; JSON; XML	RDF
PAÍS VASCO		XLS/XLSX; SHP	CSV; XML; JSON/GEOJSON	

Tabla 28. Clasificación de formatos de datos abiertos según el modelo 5 estrellas Berners-Lee por CCAA

El gráfico que viene a continuación muestra por colores el nivel de reutilización del formato en base al modelo 5 estrellas y la cantidad de CCAA que han publicado algunos de dataset en ese formato concreto. De izquierda a derecha encontraremos los formatos ordenados:

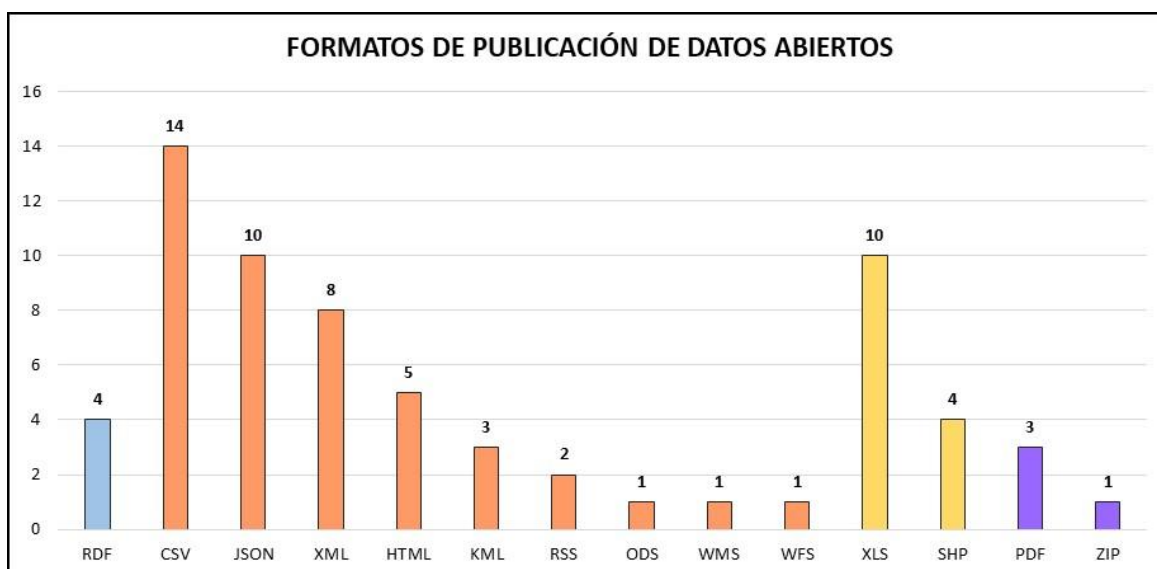


Ilustración 22. Formatos en los que publican datos de salud las CCAA

Es decir, de las dieciséis CCAA analizadas, catorce de ellas están publicando datos en CSV, equivalente en el modelo de Tim Berners-Lee a nivel 3 estrellas. Y aunque quizás sea el formato más estandarizado para publicar actualmente, no es el único formato de 3 estrellas que se está utilizando en España: JSON ha sido utilizado por diez CCAA, XML utilizado por ocho, WMS y WFS por una CCAA, y KML por 3 CCAA. El resto de formatos (a excepción de tres de ellos) pertenecen al nivel 2 estrellas. Cabe destacar también que, aunque los formatos HTML, PDF y ZIP han sido considerados como no reutilizables (por su baja posibilidad de obtener y filtrar datos) la utilización de ellos para publicar es frecuente. Tan solo 4 CCAA han publicado algún dato en RDF. Considerado como un formato 4 estrellas dado su nivel elevado de reutilización.

5.4 APLICACIÓN DE INDICADORES EXTRAIDOS DE LA NORMA UNE 178301:2015 A LOS 16 PORTALES DE DATOS ABIERTOS ANALIZADOS EN ESTE TRABAJO.

Actualmente la norma UNE 178301:2015 es considerada una norma de calidad, pero no es un documento de obligado cumplimiento por parte de las AAPP. Aun así, su uso como herramienta de trabajo aporta constancia y rigor a trabajos de investigación. Por este motivo, se ha decidido valorar una serie de indicadores de la norma (bajo criterio propio

de la investigadora y acordes a los objetivos del trabajo) y aplicarlos a los PDA autonómicos objetos de estudio en este trabajo.

Indicador de la Norma 4.2.1.2, ¿el conjunto de datos aporta información sobre condiciones de uso y licenciamiento?

Para la extracción de estos datos se ha tenido en cuenta como criterio que los conjuntos de datos ofrecidos en el portal contengan el campo de información *licencia*. En caso negativo, que el portal contenga un apartado sobre condiciones de uso o licencias en el portal que sean aplicables a todos los conjuntos de datos. Tras la comprobación pertinente de cada uno de los portales y varios dataset elegidos al azar de cada conjunto, se ha comprobado que un 13% de las CCAA no cumple con dicho criterio. Es el caso de Cataluña y Cantabria. Cabe resaltar del caso de Cataluña, que a pesar de que la web datos.gob.es facilita la información de sus licencias, el portal de esta comunidad autónoma no ofrece dicha información, o al menos de manera visible al posible usuario reutilizador. El 87% restante, tiene definido el tipo de licencia aplicada a cada conjunto de datos correspondiente, o si no, contiene un apartado visible sobre condiciones de uso.

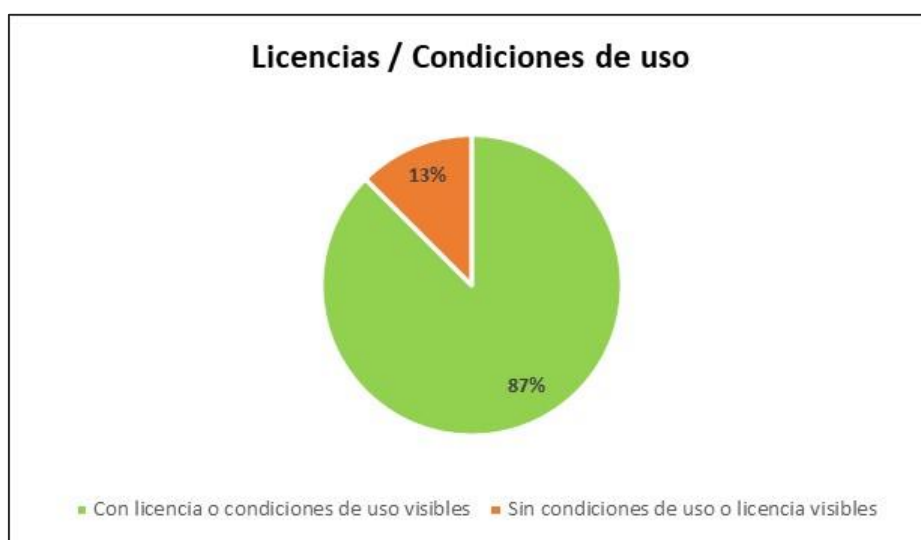


Ilustración 23. Condiciones de uso y licencias en PDA

Indicador de la Norma 4.3.2.2 b, ¿existe una medición de uso de los conjuntos de datos (descargas)?

Respecto a este indicador, se ha definido como criterio que los conjuntos de datos ofrecieran el número de descargas del mismo. Solo dos comunidades autónomas, Islas Baleares y Cataluña cumplen con este requisito. El caso de Aragón, ofrece un apartado con los dataset más descargados, pero no indica ni siquiera el número de descargas del conjunto de datos por lo que se ha descartado en este indicador.

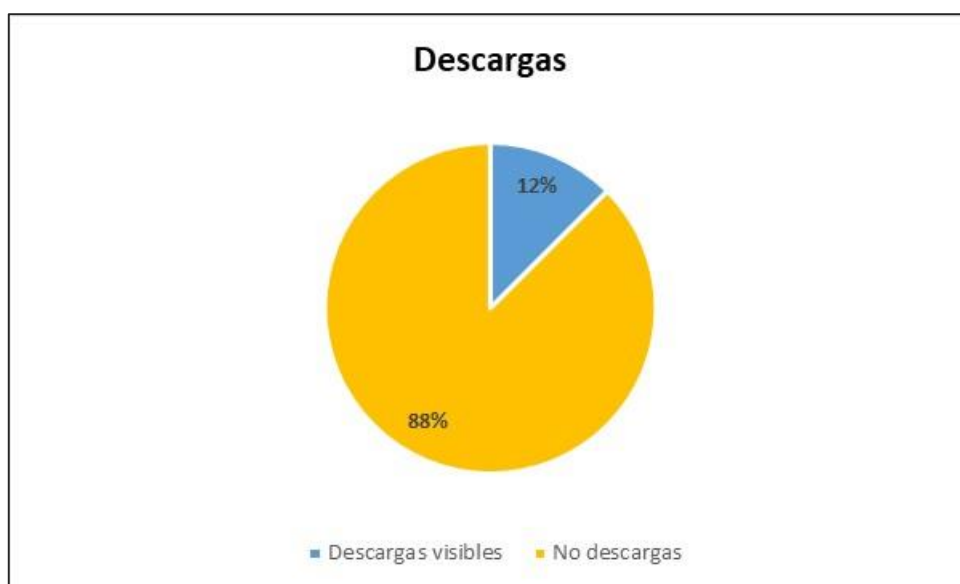


Ilustración 24. Número de descargas de conjuntos de datos

Indicador de la Norma 4.4.1.1, ¿disponen de catálogo de datos propio que proporcione el acceso a los conjuntos de datos?

Este indicador da un resultado del 100%. Su lógica es evidente. Todo portal de datos abiertos tiene como herramienta básica e indispensable un catálogo donde volcar los conjuntos de datos para ser consultados por los reutilizadores. Si bien es cierto, que algunos catálogos están más desarrollados que otros.

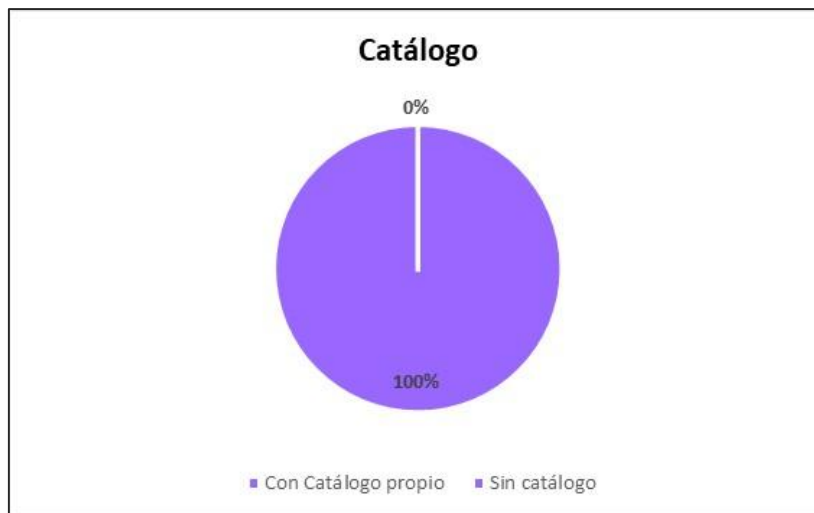


Ilustración 25. Portales de datos abiertos con catálogo de datos

Indicador de la Norma 4.4.1.3, ¿los conjuntos de datos publicados en el catálogo están documentados de acuerdo a la NTI?

Un amplio porcentaje de PDA tienen bien documentados los conjuntos de datos publicados, lo cual es un rasgo muy positivo para facilitar la reutilización de datos. En cualquier caso, hay comunidades autónomas que deberían seguir trabajando en este aspecto: facilitar de manera más sencilla la información esencial de un dataset. En el caso de Asturias, sus conjuntos de datos no tienen definidos los campos requeridos por la NTI, así que mucho menos los campos opcionales. Obtener información sobre el publicador o fecha de actualización resulta una tarea tediosa y en algunas ocasiones imposible de obtener. Los conjuntos de datos publicados en salud ofrecen información sobre el título, formato y url de acceso.

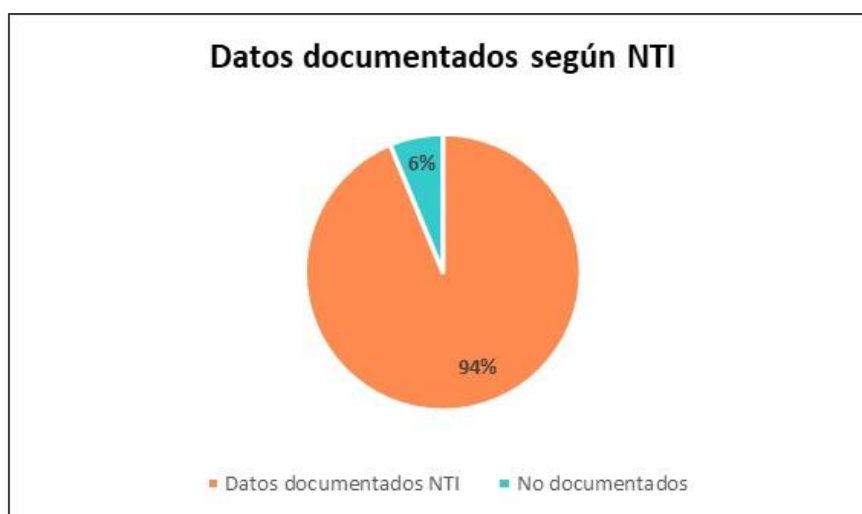


Ilustración 26. Conjuntos de datos documentados en los PDA.

Indicador de la Norma 4.4.1.4, ¿el catálogo dispone de mecanismos de búsqueda de los conjuntos de datos con los siguientes criterios: texto, formato y categorías?

Para comprender los resultados de este indicador, cabe destacar antes que, para proceso de obtención de datos de este trabajo, ha sido un punto muy positivo que el catálogo dispusiera de los mecanismos de búsqueda indicados anteriormente: texto, formato y categorías. El hecho de que dispusiera filtrados por formato o categorías ha permitido una recogida más sencilla, eficaz y rápida a la investigadora.

El 31% que muestra el gráfico, corresponde a aquellas CCAA cuyo catálogo de datos abiertos no dispone de mecanismos de búsqueda por formato o categoría. Obtener los datos de estas comunidades autónomas en cuanto a total de conjuntos de datos por formato o categorías, ha sido un trabajo laborioso.

Los PDA que no disponen de dichos mecanismos son Aragón, Asturias, Cataluña, e Islas Baleares.

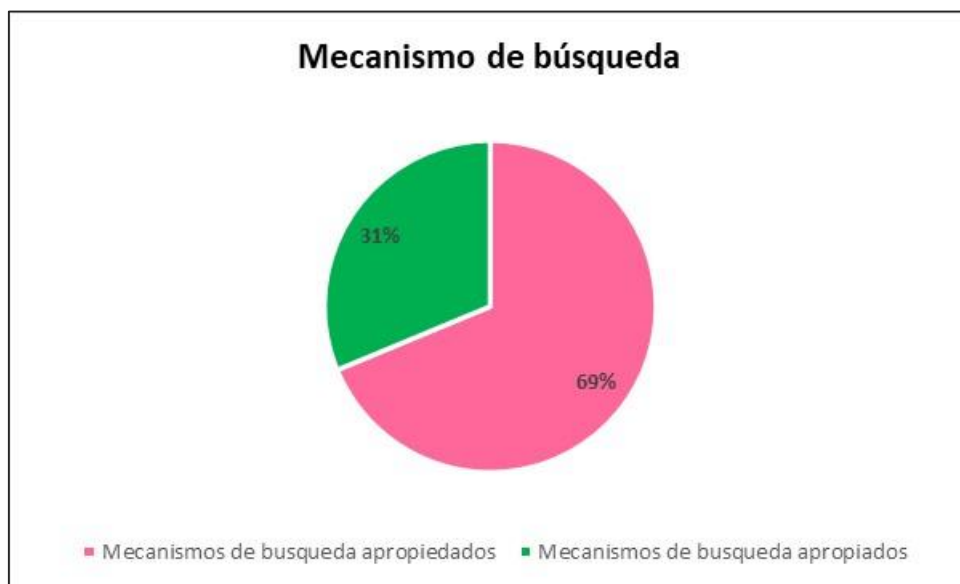


Ilustración 27. Mecanismos de búsqueda de los PDA

Indicador de la Norma 4.4.1.6, ¿el catálogo está configurado de forma que asegure que los conjuntos de datos tienen una dirección permanente en el tiempo y es fácilmente accesible?

Respecto a este indicador, el resultado ha venido determinado por la accesibilidad a los conjuntos de datos. El 19% corresponde a las CCAA de Cataluña, Islas Baleares y Murcia. En el caso de las dos primeras, el hecho de poseer un catálogo en dos idiomas ha resultado en alguna ocasión muy incómodo para acceder a los dataset, por ejemplo, cuando se pretendía acceder a un conjunto de datos en el PDA de Cataluña y el idioma lo tenía establecido en español, se cambiaba automáticamente a catalán, y viceversa. Tener que cambiar hasta dos y tres el idioma para acceder correctamente al conjunto de datos en sí terminó por resulta descartarlo en este indicador. Lo mismo ha ocurrido con Islas Baleares. En el caso de Murcia, se comprobó que las direcciones de varios dataset de salud no funcionaban correctamente.

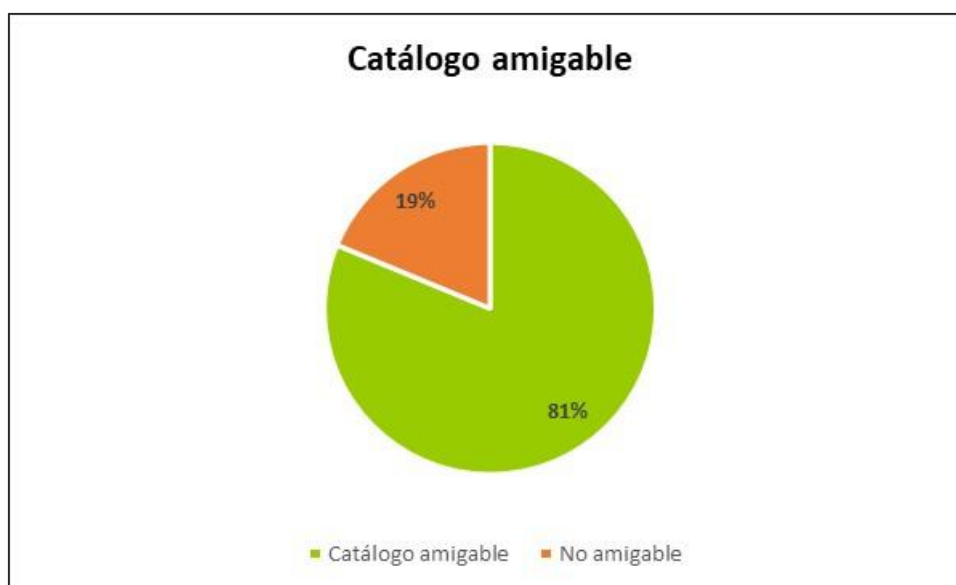


Ilustración 28. Catálogos de los PDA

Indicador de la Norma 4.4.2.3, ¿el catálogo facilita los sistemas de acceso a los datos de los conjuntos de datos para su uso con medidas API Web o un servicio SPARQL?

En el siguiente indicador, poco más de la mitad de las CCAA disponen de un sistema de acceso API Web o SPARQL. Recordemos que este componente en un portal de datos abiertos no es obligatorio, pero si recomendable. Andalucía, Aragón, Canarias, Cataluña, Islas Baleares, La Rioja y País Vasco conforman ese 56% que dispone de dicho elemento.

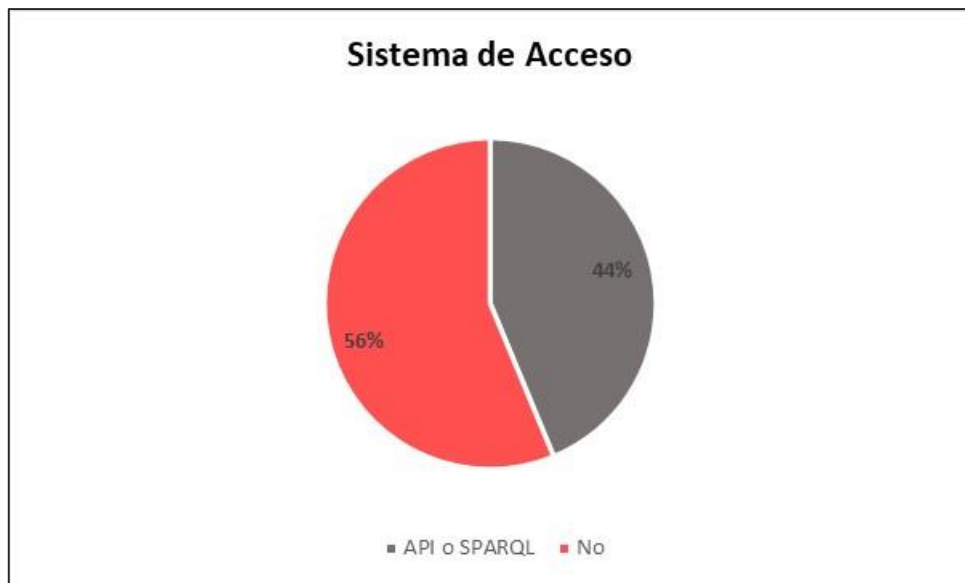


Ilustración 29. Disponibilidad de sistema de acceso API o SPARQL

Indicador de la Norma 4.5.2.1, ¿el organismo dispone de algún canal específico para enviar opiniones, participar, colaborar o pedir la publicación de datos o mejora de los actuales?

Para una mejor retroalimentación entre AAPP y usuarios reutilizadores, es importante que en los PAD exista un canal donde confluya el intercambio de información entre los participantes del portal, de cara a mejorar aspectos en relación con la apertura de datos.

Desde este punto de vista existe un 26% de CCAA que aún no disponen de un apartado que permita realizar estas acciones. Asturias y Cantabria, deberían definir en sus políticas de open data incluir en sus PDA dicho apartado.



Ilustración 30. Disponibilidad de canal de Colaboración y Participación

Indicador de la Norma 4.5.2.2, ¿el organismo facilita la resolución de quejas y conflictos sobre datos abiertos?

A diferencia del indicador anterior, donde la mayoría de las CCAA creían conveniente disponer de un canal específico de colaboración y participación, tan dos comunidades autónomas, La Rioja y Canarias, creen igual de oportuno recoger en un apartado propio la parte negativa de la opinión de un usuario para utilizarla a su favor y mejorar, como es el caso de reclamaciones, conflictos y quejas.



Ilustración 31. Canal de resolución de quejas y conflictos

Indicador 4.5.2.3, ¿la AAPP fomenta la reutilización de datos mediante recursos formativos?

Todo proyecto de apertura de datos debería ir respaldado por una serie de actividades que permitieran fomentar el conocimiento acerca de la reutilización de datos para potencial su apertura. En este sentido, varias CCAA han apostado por incluir en sus portales iniciativas de formación o educación en el uso de datos. Andalucía, Aragón, Canarias, Cataluña, Galicia y País vasco han sido las CCAA pioneras en este apartado que cobra especial relevancia de cara al futuro.



Ilustración 32. Fomento de la reutilización de datos en los PDA

A modo de resumen, se recoge en el último indicador el total de puntos que cada CCAA tiene después de haber aplicado los anteriores indicadores a cada PDA. De esta manera se comprobará que PDA cumple más requisitos según la Norma Técnica de Interoperabilidad. Para establecer el *ranking* de las CCAA se ha otorgado un punto a cada autonomía por cada ítem de información existente en el portal.

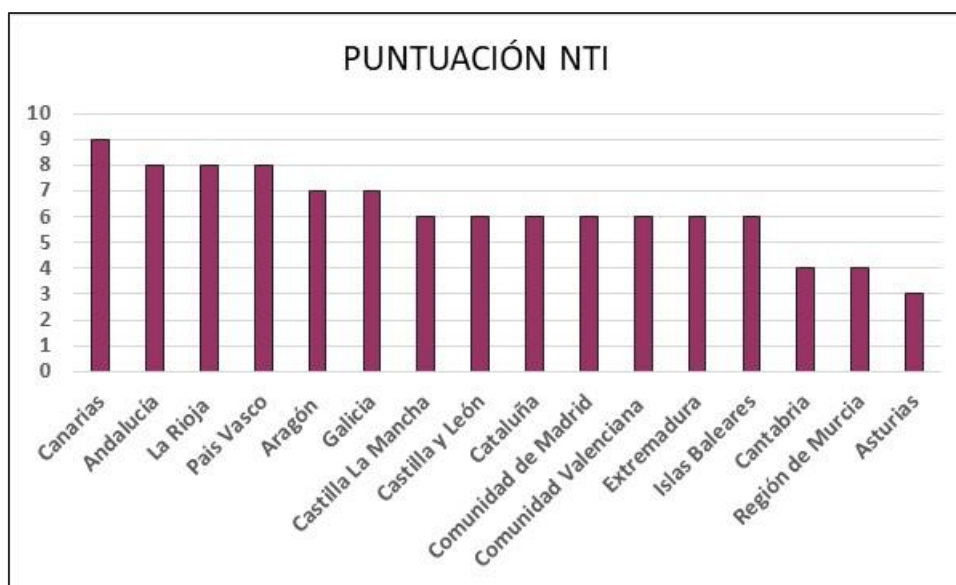


Ilustración 33. Puntuación de los PDA según los indicadores establecidos por la NTI

Según el gráfico, Canarias (9 puntos), Andalucía (8 puntos), La Rioja (8 puntos) y País Vasco (8 puntos) son los PDA más completos en torno a la NTI. Por el contrario, Asturias con 3 puntos y Murcia y Cantabria con 4 puntos respectivamente, se convierten en los PDA más incompletos; las AAPP deberían seguir trabajando en ellos para mejorarlos. Y pesar de las diferencias, un 80% de las CCAA supera el promedio de indicadores aplicados en sus portales. Este hecho permite demostrar el interés por parte de las AAPP de la corriente de apertura de datos en sus portales, al integrar en ellos apartados que permitan su acceso y reutilización.

6. CONCLUSIONES

En España, las CCAA son las principales gestoras de la mayor parte de servicios de carácter social dirigidos a los ciudadanos. Sus numerosas competencias y el elevado número de tareas que desarrollan diariamente dan como resultado la generación de un alto volumen de datos en el desarrollo de éstas. Esta notoriedad nos ha invitado a analizar cómo está siendo la divulgación de información relativa a la gestión autonómica que se está llevando a cabo, un campo de investigación apenas explotado hoy en día.

Este estudio se ha ocupado de estudiar y profundizar las distintas iniciativas de open data llevadas a cabo por las AAPP en España, de manera que pudiera ofrecerse una visión general del estado actual en publicación de datos, sobre todo en sanitarios. Pero para poder entender que se está publicando y cómo, es necesario conocer primero quienes están a favor de las iniciativas open data.

La mayoría de las AAPP autonómicas son conscientes del valor potencial que pueden tener los datos que generan, es por ello que todas (a excepción de las CCAA de Ceuta y Melilla) disponen de portales de datos abiertos. Bien es cierto, que las diferencias entre unos portales y otros son notorias. Si hacemos referencia al contenido del portal en su conjunto, disponibilidad de catálogo, aplicaciones, información de interés, apartados de participación, licencias visibles, etc, se ha podido comprobar que las CCAA con un portal más desarrollado o completo suele coincidir con los portales con más datos publicados, como es el caso del País Vasco o La Rioja. Además, se ha podido verificar que son los portales con mejor acceso a los datos, al poseer filtros de búsqueda que facilitan su acceso. El hecho de que un portal permita filtrar por formatos, ha demostrado ser una gran ventaja y muy útil, ya que permite al reutilizador acceder de forma directa a aquellos conjuntos de datos fácilmente legibles por máquinas.

Centrándonos en el catálogo de datos encontramos fallos generales. Un error por parte de ciertas administraciones es disponer de buscadores de datos excesivamente acotados o guiados. Los buscadores excesivamente acotados generalmente no son suficientes para garantizar el acceso a todos los datos, independientemente si contamos con más o menos opciones de filtrado. Mientras que los buscadores excesivamente guiados no han concedido suficiente libertad a la investigadora para acceder a los datos según sus necesidades. Este fue el caso del PDA de Navarra que, dada la complejidad para garantizar el acceso real a datos sanitarios, se decidió descartar en el análisis de este estudio.

Respecto los conjuntos de datos publicados y las temáticas, los PDA recogen registros de todo tipo (información medioambiental, información económica, social, cultural, legislativa, etc.). Es un poco acierto por parte de ciertas AAPP, publicar las categorías de datos en el catálogo siguiendo sus propios criterios. La NTI no indica que se deba utilizar la taxonomía de temas expuesta en ésta para organizar la información internamente, o para los buscadores, pero esto permite clasificar el catálogo y los conjuntos de datos de forma común usando los mismos mecanismos, éste hecho facilitaría aun más el acceso a los datos a nivel de todos los portales. En cuanto a la diversidad de categorías definidas puede estar motivada por las distintas opciones de creación de valor que se les atribuye a los datos abiertos por parte de las AAPP. Según los resultados mostrados en el apartado anterior, abunda más la información de tipo medio ambiental, económica o relacionada con el sector público; sectores de la sociedad cuyos datos pueden generar más economía a una administración a corto plazo, y sobre todo porque estos datos pueden ser recopilados de forma más rápida y sencilla por parte de las administraciones. El análisis de este trabajo, lo corrobora. La diferencia que se establece a nivel general entre la publicación de conjuntos de datos en medio ambiente, sector público, economía o demografía es demasiado significativa con respecto a la publicada en salud. En algunos casos supera la diferencia de mil dataset.

En esta misma línea, todos los portales de datos autonómicos coinciden en el hecho de recoger datos en salud (dado que es un sector que preocupa a la sociedad), y es un gran avance, pero las cantidades de dataset recogidos por las administraciones son muy escasos. Tras la revisión de la literatura para la realización de este trabajo, se cree que los motivos pueden venir determinados por la dificultad de consenso a la hora publicar y la falta de interoperabilidad entre sistemas públicos de salud, por las restricciones y estricta privacidad de datos que existe actualmente o precisamente por la falta de un modelo sanitario al que seguir. Aun así, comunidades como Aragón o País Vasco hacen mayores esfuerzos por ofrecer datos en materia de salud, y comunidades como Asturias o Canarias deberían tomar como referente nacional las políticas de apertura de éstas para aumentar la publicación de sus datos.

Esto nos invita a hacer la siguiente reflexión ¿el alto porcentaje de datos publicados por un organismo es directamente proporcional a la calidad de esos datos? La respuesta es no, o al menos no es concluyente. No ha sido posible afirmar qué porque una CCAA tenga más datos publicados la calidad de éstos va a ser más alta. El análisis ha mostrado que aquellas CCAA que usan más formatos estructurados para la publicación de sus datos han sido Cataluña y Galicia, comunidades autónomas que tienen un total de 23 conjuntos de datos. Se deduce de este resultado que en función del tipo de información que se quiere publicar se utilizará un tipo de formato u otro (al mismo nivel de calidad) pero eso no hace menos válido los conjuntos de datos publicados por AAPP en un único formato estructurado y no propietario, aunque es aconsejable adaptar los formatos. Por ejemplo, utilizar JSON para datos georeferenciados y CSV para datos estadísticos. Además, la ausencia de filtrado por formatos en algunos portales, ha derivado en el recuento manual de los dataset por parte de la investigadora, lo que ha sido un trabajo demasiado laborioso si tenemos en cuenta que uno de los principios de los datos abiertos es su fácil acceso

Lo que sí ha quedado en evidencia es que la publicación de datos en formatos no reutilizables es una práctica que deberían ir reduciendo las AAPP, a favor de la utilización de formatos más avanzados para la apertura de datos. Aunque el hecho de que la gran mayoría de las iniciativas identificadas hayan facilitado reutilización de los datos en formatos de 3 estrellas, pone de manifiesto el interés de los organismos publicadores por seguir impulsando la reutilización de sus datos abiertos. La Comunidad de Madrid es un buen ejemplo, ya que las publicaciones de sus datos en salud solo han sido divulgadas en este nivel.

En definitiva, se puede afirmar que en líneas generales España es propulsora de la reutilización de información en el sector público al contar con numerosas iniciativas de open data; determinar qué posición ocupa con respecto a Europa es otra cuestión por investigar, pero no ocupa lugar en este trabajo. Los aspectos que deberían mejorar las administraciones públicas respecto a sus portales y calidad de datos, pasa simplemente por utilizar como referencia en sus trabajos las normas vigentes en materia de apertura de datos, como es el caso de LRIPS o la NTI. La estandarización de todos los portales en base a estas normas permitiría caminar en una misma dirección: facilitar el acceso a los datos públicos.

Respecto a la publicación de datos en salud (objetivo de análisis en esta investigación) las comunidades autónomas españolas tienen trabajo que hacer todavía. La implicación por parte de los gobiernos se vuelve fundamental para llegar a un consenso nacional que permita redactar una política de apertura de datos sanitarios común y aumentar no solo la publicación de datos por parte de las administraciones, si no fomentar la utilización de formatos de calidad. Las AAPP también tendrán que seguir poniendo todos sus esfuerzos en publicar datos de calidad, ya que puede existir una tendencia por parte de las administraciones a publicar datos solo por ver engrosado el número de dataset publicados, sin valorar realmente si la información que ofrecen puede ser de utilidad o no.

De cara a un futuro, el mayor reto en materia de open data en sanidad se presentará en la transformación de los datos no estructurados (recetas de papel, los registros médicos, las notas manuscritas de médicos y enfermeras, las grabaciones de voz, las radiografías, escáneres, resonancias magnéticas, TAC, otras imágenes médicas, etc) en datos

estructurados para organizarlos y definirlos de manera que puedan utilizarse para predecir, prevenir y personalizar enfermedades, definir causas de mortalidad, ofrecer mejores soluciones médicas, etc. Además, esta transformación podría ayudar también a la contabilidad electrónica, gestión administrativa (ineficiencias administrativas), manipulación de datos clínicos (respetando la privacidad de los datos del paciente), mejoras en la coordinación de atención al paciente, lucha contra abusos, etc.

7. BIBLIOGRAFÍA

- Aberasturi Gorriño, U. (2013). La protección de datos en la sanidad, Thomson Reuters-Aranzadi. Cizur Menor, Navarra: Aranzadi, 2013, pp. 73 y ss.
- Almansa Morales, S. (2017). Transparencia y datos abiertos en la administración pública. Madrid: Instituto Nacional de Administración Pública (INAP).
- Andreu Martinez, A. (2017). Open Data en el ámbito sanitario y su compatibilidad con la privacidad del paciente. *Revista Internacional de Gobiernos Abiertos*, v. 5. Recuperado de http://ojs.imodev.org/index.php/RIGO/article/view/200/330#_ftn1.
- Armstrong, E. (2005). Integrity, Transparency and Accountability in Public Administration: Recent Trends, Regional and International Developments and Emerging Issues. United Nations: Department of Economics and Social Affairs. Recuperado de http://digibuo.uniovi.es/dspace/bitstream/10651/38527/6/TFM_HuergoRodriguezAlejandra.pdf
- Arquero Avilés, R. y Marco Cuenca, G. (2014). El Portal de datos abiertos de la Unión Europea: análisis y evaluación. *Revista General De Información Y Documentación*, 24(1), 99-118. Recuperado de https://doi.org/10.5209/rev_RGID.2014.v24.n1.45384
- Atresmedia (2015) Las combinaciones en las licencias creative Commons. Recuperado de <http://www.blogcreacultura.com/combinaciones-creative-commons/>
- Centro de Estudios de Tecnologías Web (s.f). Guía de datos abiertos. Recuperado de <https://ceweb.br/guias/dados-abertos/es/capitulo-4/>
- Cerrillo, A. (2014). Los principios de los datos abiertos en la legislación española. *Revista de Internet, Derecho y Política*. Núm. 19.
- Comisión Europea (2011). CComunicación de la comisión al parlamento europeo, al consejo, al comité económico y social europeo y al comité de las regiones. Datos abiertos: un motor para la innovación, el crecimiento y la gobernanza transparente. COM (2011) 882 final. Recuperado en <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52011DC0882&from=ES>

- Curto Rodríguez, R. (2017). Datos Abiertos y rendición de cuentas en las comunidades autónomas españolas. Niveles de divulgación y determinantes (tesis doctoral) Universidad de Oviedo, Departamento de contabilidad. Oviedo
- Dorado Pérez, E. (2019). "Portales de datos abiertos. Experiencia práctica en el Consejo de Transparencia y Protección de datos en Andalucía". *Revista española de transparencia*, v. 2, n. 2, 33-38
- European Commission. Programa de formación en línea del Portal Europeo de Datos. Recuperado en <https://www.europeandataportal.eu/elearning/es/#/id/co-01>
- Federación Española de Municipios y Provincias (2017). Datos Abiertos: guía estratégica para su puesta en marcha. Conjunto de datos mínimos a publicar. Recuperado de <http://femp.femp.es/files/3580-1617-fichero/Gu%C3%ADa%20Datos%20Abiertos.pdf>.
- Fernandez Salmerón, M. (ed.). (2014). *Régimen jurídico de la transparencia del sector público: del derecho de acceso a la reutilización de la información*. Cizur Menor: Thomson-Aranzadi, pág. 376.
- Ferrer-Sapena, A.; Peset, F.; Aleixandre- Benavent, R. (2011). Acceso a los datos públicos y su reutilización: open data y open government. *El profesional de la información*, v. 20, n. 3, 260-269.
- García García, J. (2010). *Gobierno digital y responsabilidad pública en los municipios españoles*. (Tesis Doctoral) Universidad de Oviedo, Departamento de Contabilidad.
- García García, J. y Curto Rodriguez, R. (2019). El ejercicio de la rendición de cuentas mediante portales de datos abiertos en las comunidades autónomas españolas. IDP, n.º 29. Recuperado de <https://dialnet.unirioja.es/servlet/tesis?codigo=176185>
- García Saura, P. J. (2014). La reutilización de la información en el sector turístico. Acceso por parte del entramado empresarial a la información administrativa y su posterior reutilización comercial. *Cuadernos de turismo*, 34, 89-113. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=4764259>
- Garriga-Portolà, M. (2011). ¿Datos abiertos? Sí, pero de forma sostenible. *El profesional de la información*, v. 20, n. 3, pp. 298-303. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=3679618>

- Generalitat de Catalunya (s.f). Gobierno Abierto: ¿qué beneficios tienen los datos abiertos? Recuperado de http://governobert.gencat.cat/ca/dades_obertes/que-son-les-dades-obertes/
- Generalitat de Catalunya (2017). Programa de analítica de datos para la investigación y la innovación en salud (PADRIS). Recuperado de: http://aquas.gencat.cat/web/.content/minisite/aquas/publicacions/2017/Programa_analitica_dades_PADRIS_aquas2017.pdf
- Gértrudix, M.; Gertrudis-Casado, M.C.; Álvarez- García, S. (2016). Consumption of public institutions' open data by Spanish citizens. *El profesional de la información*, v. 25, n. 4, pp. 535-544. <https://doi.org/10.3145/epi.2016.jul.03>
- Janssen, M.; Charalabidis, Y.; Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information systems management*, v. 29, n. 4, pp. 258-268.
- Llinares, J. (2010). Las 10 claves para entender la colaboración en el Modelo Open Government. En: C. L. CALDERÓN, SEBASTIÁN (ed.) Open Government: Gobierno Abierto, Alcalá la Real: Algón, 2010
- Martín Delgado, I. (2014). Transparencia, reutilización y datos abiertos. Algunas reflexiones generales sobre el acceso libre a la información pública. En: J. Valero Torrijos y M. Fernández Salmerón (ed.). Régimen jurídico de la transparencia del sector público: del derecho de acceso a la reutilización de la información. Cizur Menor: Thomson-Aranzadi, pág. 376.
- Martín González, Y. (2013). Política europea de reutilización de la información del sector público. De la norma jurídica al portal de datos abiertos. *RUE: Revista universitaria europea*, 19, 113-134.
- Maseda-Seco, D.; Bueno-de-la-Fuente, G.; Méndez, E. (2017). Análisis y categorización de los datos abiertos de las bibliotecas municipales españolas: metadatos, interoperabilidad y propuesta para la apertura y reutilización. *El profesional de la información*, v. 26, n. 3, pp. 392-402. <https://doi.org/10.3145/epi.2017.may.05>
- Moore, M. (2018). Europa ha desarrollado unas políticas sólidas de Open Data, pero debe seguir implementando medidas estratégicas y promoviendo la concienciación. Capgemini.

Recuperado de <https://www.capgemini.com/es-es/wp-content/uploads/sites/16/2018/11/NP-Informe-Open-Data.pdf>

Obama, B. (2009). "Transparency and open government. Memorandum for the heads of executive departments and agencies", en *Federal Register*, nº 15, vol. 74, pp. 4685–4686.

Open Government Partnership (2011). Plan de gobierno abierto. Recuperado de <http://www.opengovpartnership.org>

Open Knowledge Foundation (s.f) Open data handbook. ¿Qué son los datos abiertos? Reino Unido. Recuperado de <http://opendatahandbook.org/guide/es/what-is-open-data/>

Oviedo Blanco, E. (2016). Modelo de madurez para portales de datos abiertos e incorporación a la norma técnica nacional de Costa Rica (tesis doctoral). Universidad de Alicante, Alicante, España. Recuperado de https://rua.ua.es/dspace/bitstream/10045/54316/1/tesis_edgar_oviedo_blanco.pdf

Parlamento Europeo (2013), "Directiva 2013/37/ue del Parlamento Europeo y del Consejo, de 26 de junio de 2013 por la que se modifica la Directiva 2003/98/ce, relativa a la reutilización de la información del sector público," en *Diario Oficial de la Unión Europea* (27/06/2013). Recuperado de <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32013L0037&from=ES>

Pintos Teigeiro, C.; Marín de la Iglesia, J.L (2013). Open data. Reutilización de la información pública. Madrid: Instituto Nacional de Administración Pública (INAP): Netbiblo.

Pintos, C (2012.) Open data. Reutiliza la información pública para crear una nueva empresa. Madrid: Netbiblo, 2012.

Portal de datos abiertos de Madrid. Formatos. Recuperado de <https://datos.madrid.es/portal/site/egob/>

Ramos Simón, L.F., Arquero Avilés, R., Botezan, I., Cobo Serrano, S., Sala Jiménez, A., Sala Jiménez, R., & Del Valle Gastamiza, F. (2012). De la reutilización de información del sector público a los portales de datos abiertos en Europa, *BiD: textos universitarios de biblioteconomía i documentació* 29. Recuperado de <http://bid.ub.edu/29/pdf/ramos2.pdf> [Fecha de consulta 20/10/2019]

- Rovira, M. (s.f), Las TIC en la coordinación entre los sistemas social y sanitario, Informática+Salud, nº 116, 2016, p. 6. Recuperado de <http://www.pmfarma.es/noticias/22452-documento-para-impulsar-la-transformacion-digital-del-sector-sanitario.html>
- Serrano Perez, M.M. (2015). Big data o la acumulación masiva de datos sanitarios: derechos en riesgo en el marco de la sociedad digital. *Revista Derecho y Salud*, nº 25, 2015, pp. 51 y ss.
- Sunlight Foundation (s.f). 10 principles of Open Government Data. Recuperado de <https://sunlightfoundation.com/policy/documents/ten-open-data-%20principles/>
- Unión Europea. Open Government Data Principles. Recuperado de https://public.resource.org/8_principles.html
- Valero Torrijos (2014). “Acceso, reutilización y gestión avanzada de la información sanitaria en el ámbito de la Administración sanitaria: implicaciones jurídicas desde la perspectiva de la innovación tecnológica”, en J. Valero Torrijos y M. Fernández Salmerón, Régimen jurídico de la transparencia del sector público. Del derecho de acceso a la reutilización de la información, Thomson Reuters-Aranzadi, Cizur Menor (Navarra), 2014, pp. 643 y ss.
- Yu, H., Robinson, D.G. (2012). «The New Ambiguity of “Open Government”». *UCLA Law Review*, N.º 59, págs. 178-208. <https://doi.org/10.2139/ssrn.2012489>

NORMATIVA

- AENOR. Ciudades Inteligentes. Datos Abiertos (open data). UNE 178301. Madrid: AENOR, 2014.
- Comunicación de la Comisión Europea (2006). Libro Verde sobre la iniciativa europea a favor de la transparencia COM (2006) 194 final
- Directiva 2003/98/CE del Parlamento Europeo y del Consejo de 17 de noviembre de 2003 relativa a la reutilización de la información del sector público. Boletín Oficial del Estado, 345, de 31 de diciembre de 2003, 90 a 96. Recuperado de <https://www.boe.es/doue/2003/345/L00090-00096.pdf>

Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013, por la que se modifica la Directiva 2003/98/CE relativa a la reutilización de la información del sector público. Boletín Oficial del Estado, 175, de 27 de junio de 2013, 1 a 8. Recuperado de <https://www.boe.es/doue/2013/175/L00001-00008.pdf>

Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. Transpone la Directiva 2003/98/CE, de 17 de noviembre de 2003. Boletín Oficial del Estado, 276, de 17 de noviembre de 2007, 47160 a 47165. Recuperado de <https://www.boe.es/boe/dias/2007/11/17/pdfs/A47160-47165.pdf>

Ley 18/2015, de 9 de julio, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. Boletín Oficial del Estado, 164, de 10 de julio de 2015, 57436 a 57450. Recuperado de <https://www.boe.es/boe/dias/2015/07/10/pdfs/BOE-A-2015-7731.pdf>

Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno. Boletín Oficial del Estado, 295, de 10 de diciembre de 2013, 97922 a 97952. Recuperado de <https://www.boe.es/boe/dias/2013/12/10/pdfs/BOE-A-2013-12887.pdf>

Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Boletín Oficial del Estado, 295, de 10 de diciembre de 2013, 97922 a 97952. Recuperado de <https://www.boe.es/boe/dias/2013/12/10/pdfs/BOE-A-2013-12887.pdf>

Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica. Boletín Oficial del Estado, 274, de 15 de noviembre de 2002, 40126 a 40132. Recuperado de <https://www.boe.es/boe/dias/2002/11/15/pdfs/A40126-40132.pdf>

Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal. Boletín Oficial del Estado, 269, de 8 de noviembre de 2011, 116296 a 116307. Recuperado de <https://www.boe.es/boe/dias/2011/11/08/pdfs/BOE-A-2011-17560.pdf>

Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica. Boletín Oficial del Estado, 25, de 29 de enero de 2010, 8139 a 8156. Recuperado de <https://www.boe.es/boe/dias/2010/01/29/pdfs/BOE-A-2010-1331.pdf>

PORTALES DE DATOS ABIERTOS DE LAS COMUNIDADES AUTÓNOMAS

Portal de Datos Abiertos Junta de Andalucía. Junta de Andalucía
<https://www.juntadeandalucia.es/datosabiertos/portal.html>

Aragón Open Data. Gobierno de Aragón <https://opendata.aragon.es/>

Open Data del Principado de Asturias. Principado de Asturias.
<https://www.asturias.es/portal/site/webasturias/menuitem.7b2ff7592927f53571a70084ec12b2a0/?vgnextoid=2cab7cd61f918510VgnVCM100000ce212b0aRCRD&i18n.http.lang=es>

Open Data GOIB. Gobierno de las Islas Baleares
<http://www.caib.es/sites/opendatacaib/es/inicio-27248/?campa=yes>

Open Data del Gobierno de Canarias. Gobierno de Canarias
<https://opendata.gobiernodecanarias.org/opendata/>

Datos Abiertos del ICANE. Instituto Cántabro de Estadística <https://datos.icane.es/>

Datos Abiertos CLM. Junta de Castilla La Mancha <https://datosabiertos.castillalamancha.es/>

Portal de Datos Abiertos Junta de Castilla y León <https://datosabiertos.jcyl.es/web/es/datos-abiertos-castilla-leon.html>

Portal Datos Abiertos Generalitat de Catalunya. Generalitat de Catalunya
http://governobert.gencat.cat/ca/dades_obertes/

Datos Abiertos de la Generalitat Valenciana. Gobierno de Valencia
http://www.dadesobertes.gva.es/va/dataset?q=&sort=metadata_modified+desc

Portal de Transparencia y Participación ciudadana de Extremadura. Gobierno de Extremadura
<http://gobiernoabierto.juntaex.es/>

Portal Open Data de la Xunta de Galicia. Xunta de Galicia <https://abertos.xunta.gal/portada>

Datos Abiertos Rioja. Gobierno de La Rioja <https://www.larioja.org/tecnologia-comunica/es/proyectos-dgtic/open-data>

Portal Datos Abiertos de la Comunidad de Madrid. Comunidad de Madrid
<http://www.comunidad.madrid/gobierno/datos-abiertos>

Datos Abiertos en la Región de Murcia. Región de Murcia
<http://datosabiertos.regiondemurcia.es/>

Portal de Gobierno Abierto de la Comunidad Foral de Navarra. Gobierno de Navarra
<http://www.gobiernoabierto.navarra.es/es/open-data>

Open Data Euskadi. Gobierno del País Vasco <http://opendata.euskadi.eus/inicio/>

ANEXO I. EVOLUCIÓN DEL LINKED GOBERNMENT DATA

20/01/1999	Comunicación de la Comisión Europea (Libro Verde) sobre la información del sector público en la sociedad de la información <i>Identifica la información del sector público clave en Europa para el desarrollo económico, el empleo y la transparencia.</i>
03/12/2001	Reglamento 1049/2001 del Parlamento Europeo y del Consejo de 30/05/2001 relativo al acceso del público a los documentos del Parlamento Europeo, del Consejo y de la Comisión. <i>Permite un amplio acceso a los documentos del Parlamento Europeo, del Consejo y de la Comisión.</i>
17/11/2003	Directiva 2003/98/CE del Parlamento Europeo y del Consejo relativa a la Reutilización de la Información del Sector Público <i>Es el primer marco comunitario regulador del proceso de apertura de archivos de los estados miembros.</i>
03/05/2006	Comunicación de la Comisión de las Comunidades Europeas (Libro Verde) sobre la iniciativa europea a favor de la transparencia. <i>Recuerda el compromiso de facilitar la participación activa en la configuración de la política de la UE y que la transparencia forma parte de la legitimidad de toda administración pública.</i>
14/03/2007	Directiva 2007/2/CE del Parlamento Europeo y del Consejo por la que se establece una infraestructura de información espacial en la Comunidad Europea (INSPIRE) <i>Fija las normas para ofrecer una información geográfica de calidad en el ámbito europeo.</i>
22/06/2007	Ley 11/2007, de acceso electrónico de los ciudadanos a los servicios públicos <i>Obliga a las entidades públicas crear canales de comunicación no presenciales con los ciudadanos. Nace el derecho a utilizar los medios electrónicos para cualquier gestión o transacción con las AA. PP.</i>
16/11/2007	Ley 37/2007, de 16 de noviembre, sobre la reutilización de la información del sector público <i>Contiene el régimen jurídico aplicable a la reutilización de los documentos en manos del sector público.</i>
08/12/2007	Los ocho principios <i>Open Government Data</i> y su evolución <i>Recomendaciones que deben cumplir los datos gubernamentales para ser considerados abiertos y de calidad a efectos de reutilización</i>
2009	Memorando Obama sobre Transparencia y GA. Directiva sobre GA <i>Los pilares del GA son la transparencia, la colaboración y la participación. La información debe ser libre por defecto y tiene que publicarse en formatos abiertos.</i>
12/03/2009	Lanzamiento del proyecto aporta.es <i>Impulsa, a nivel español, la reutilización de la información del sector público. Apertura de www.aporta.es</i>
21/05/2009	Lanzamiento del portal Data.gov <i>Portal de datos abiertos estadounidense que sirvió de ejemplo para la apertura de otros portales nacionales</i>
15/06/2009	Publicación de la guía Aporta sobre la Reutilización del sector público <i>Orientaciones a las AA. PP españolas para facilitar e impulsar la reutilización de la información</i>
30/09/2009	Lanzamiento en Reino Unido de data.gov.uk <i>Apertura del primer portal de datos abiertos nacional a nivel europeo</i>
29/01/2010	Real Decreto 4/2010 que regula el esquema nacional de interoperabilidad en el ámbito de la administración electrónica <i>Garantiza que se puedan intercambiar datos que previamente se han homogeneizado para poder ser combinables entre la misma o diferentes AA. PP.</i>
25/05/2010	Tim Berners-Lee presenta el modelo <i>5 stars open data</i> en la Expo gov2.0 celebrada en Washinton DC <i>Clasifica distintos niveles de los datos cara a su reutilización</i>
16/09/2011	Creación de la Alianza para el Gobierno Abierto (<i>Open Government Partnership</i>) <i>Asociación que promueve la transparencia y el suministro de información pública en formatos abiertos</i>

24/10/2011	Real Decreto 1495/2011 de desarrollo de la Ley 37/2007. Lanzamiento del portal datos.gob.es . (evolución del proyecto Aporta)
<i>El Real Decreto pretende facilitar al máximo la puesta a disposición de la información del sector público. El portal www.datos.gob.es, se configura como el portal de carácter nacional que organiza y gestiona el Catálogo de Información Pública de la Administración General del Estado</i>	
17/04/2012	España se une al <i>Open Government Partnership</i> y publica el “Plan de Acción OGP”. Se analiza también el II Plan de Acción
<i>Incorporación a la alianza y avances realizados en su primer y segundo plan de acción</i>	
05/09/2012	Publicación por la <i>Web Foundation</i> del modelo de medición “open data index o web index data”, una evolución del “web index”
<i>Evalúa el nivel de apertura de información en 122 países del mundo</i>	
04/03/2013	Aprobación de la Norma Técnica de Interoperabilidad de reutilización de recursos de la información
<i>Pautas básicas para la reutilización de documentos del sector público</i>	
26/06/2013	<i>Directiva 2013/37/UE del parlamento Europeo y del Consejo</i>
<i>Adaptar la anterior directiva a los cambios ocurridos en esos 10 años tanto en el volumen de información como en el progreso de las tecnologías</i>	
09/12/2013	Ley 19/2013, de transparencia, acceso a la información pública y buen gobierno
<i>Ley que busca incrementar la transparencia en la actividad pública (obligaciones de publicidad activa para la Administración), garantizar el derecho al acceso a la información y, establecer para los responsables públicos obligaciones de buen gobierno.</i>	
10/12/2014	Lanzamiento del portal de transparencia del estado
<i>Repositorio donde se deposita diversa información relacionada con la transparencia y la rendición de cuentas.</i>	
09/07/2015	Ley 18/2015, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público.
<i>Transpone la directiva 2013/37/UE al ordenamiento jurídico español</i>	

Tabla 29. Evolución del Linked Open Data.

Fuente: Curto, 2017

ANEXO II. EJEMPLO DE DATASET CON LOS CAMPOS REQUERIDOS DE INFORMACIÓN POR LA NTI

TÍTULO

Personas de 16 a 64 años en relación con la discapacidad y la actividad, y su distribución. Años 2014 a 2017.

Información del conjunto de datos

Título
Personas de 16 a 64 años en relación con la discapacidad y la actividad, y su distribución. Años 2014 a 2017.

Descripción
Recopilación de la serie 2009-2014 publicadas por el INE posibilitando la comparación de los datos referentes a la Comunidad Autónoma de Aragón y los de España.

TEMÁTICA

Categoría
Salud

Etiquetas y palabras clave
Dato estadístico Discapacidad Empleo

DESCRIPCIÓN

Organización responsable de la publicación
Instituto Aragonés de Estadística

PUBLICADOR

Territorio en el que existen datos
Este conjunto de datos contiene datos en el territorio de Otro
Esta es la URI del recurso Aragón y España

Periodo de tiempo en el que existen datos
2014 - 2017

Licencia
Creative Commons Attribution 4.0

Nivel de detalle
Hogares, personas de 16 a 64 años en relación con el empleo y la discapacidad, personas con certificado de discapacidad. Grado de discapacidad, sexo, grupo de edad, tipo de discapacidad y grado de deficiencia.

Descargas

Archivos con los datos
Personas de 16 a 64 años en relación con la discapacidad y la actividad, y su distribución. Años 2014 a 2017.

FORMATOS

px CSV

Metadatos de la colección
RDF